

Europäisches Patentamt
European Patent Office
Office européen des brevets



(11)

EP 0 955 382 A2

(12)

EUROPEAN PATENT APPLICATION

(43) Date of publication:
10.11.1999 Bulletin 1999/45

(51) Int Cl.⁶: C12Q 1/68, C07K 14/72,
G06F 17/30

(21) Application number: 99250150.2

(22) Date of filing: 07.05.1999

(84) Designated Contracting States:
AT BE CH CY DE DK ES FI FR GB GR IE IT LI LU
MC NL PT SE
Designated Extension States:
AL LT LV MK RO SI

(30) Priority: 07.05.1998 US 84641 P
03.05.1999 US 304232

(71) Applicants:
• Affymetrix, Inc.
Santa Clara, CA 95051 (US)
• Case Western Reserve University
Cleveland, Ohio 44106 (US)

(72) Inventors:
• Fan, Jian Bing
Palo Alto, CA 94306 (US)
• Chakravarti, Aravinda
Shaker Heights, Ohio 44120 (US)
• Haluska, Mark Kenneth
Cleveland Heights, Ohio 44106 (US)

(74) Representative: UEXKÜLL & STOLBERG
Patentanwälte
Beselerstrasse 4
22607 Hamburg (DE)

(54) Polymorphisms associated with hypertension

(57) The invention discloses a collection of polymorphic sites in genes known or suspected to have a role in hypertension. The invention provides nucleic acids including such polymorphic sites. The nucleic acids can

be used as probes or primers or for expressing variant proteins. The invention also provides methods of analyzing the polymorphic forms occupying the polymorphic sites.

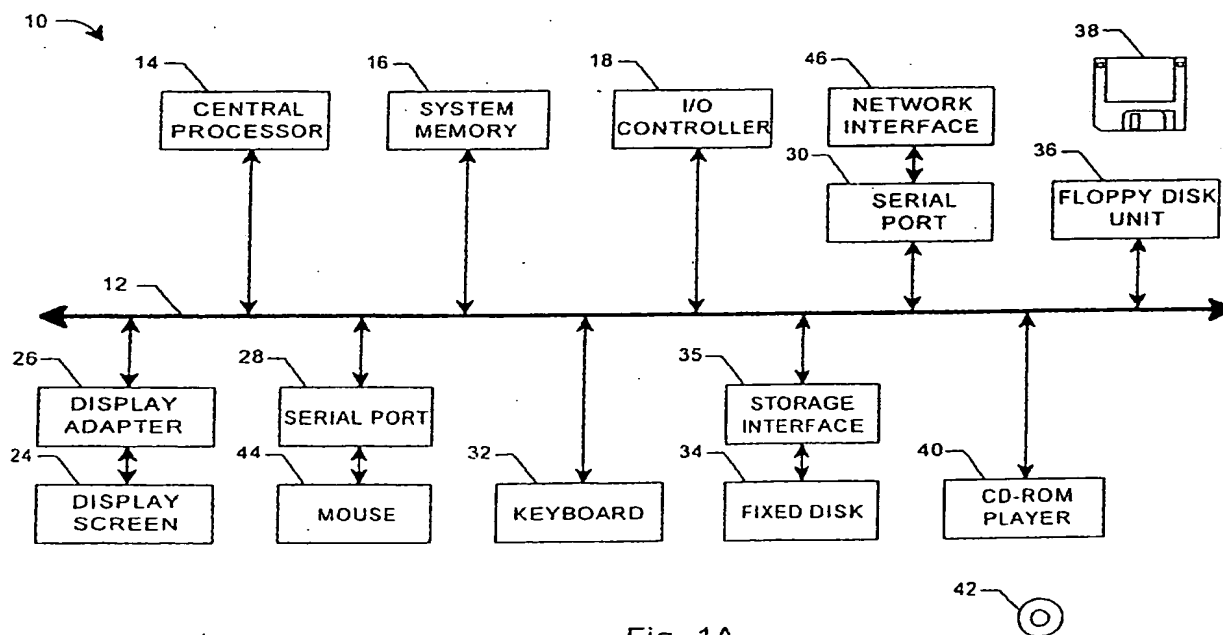


Fig. 1A

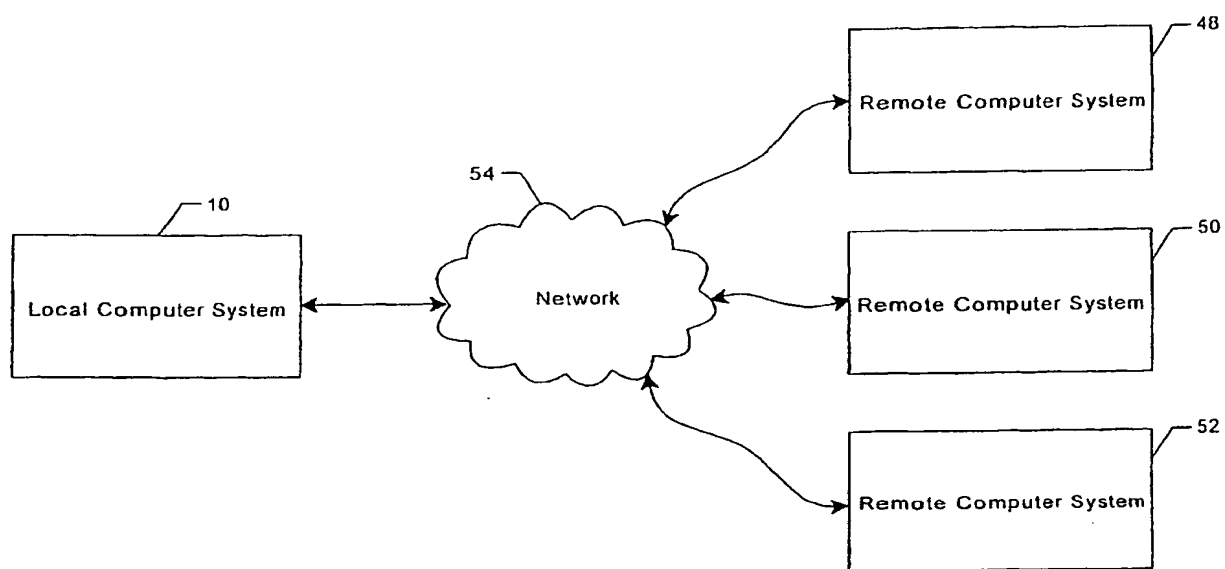


Fig. 1B

Description

CROSS-REFERENCES TO RELATED APPLICATIONS

5 [0001] This application derives priority from USSN 60/084,641 filed May 7, 1998, which is incorporated by reference in its entirety for all purposes.

[0002] The work described in this application was funded, in part, by a grant from the National Heart, Lung & Blood Institute (U10 HL54466), which may have certain rights in this invention.

10 BACKGROUND OF THE INVENTION

[0003] Hypertension, or high blood pressure, is a common disease affecting 50 million Americans and contributing to over 200,000 deaths annually from stroke, myocardial infarction, and end-stage renal disease. The disease is multifactorial and numerous genetic and nongenetic components, such as salt intake, age, diet, and body mass, are suspected to contribute. A specific cause of hypertension can typically be identified in only a small percentage of patients. Other patients with abnormally high blood pressure of unknown cause are said to have essential hypertension.

15 [0004] The existence of a genetic component to hypertension is known from twin studies, which have revealed a greater concordance of blood pressure in monozygotic twins than in dizygotic twins. Similarly, biological siblings have show greater concordance of blood pressure than adoptive siblings raised in the same household. Such studies have suggested that up to about 40% of the variations in blood pressure in the population are genetically determined.

[0005] There is a substantial pool of candidate genes that may contribute to the genetic component of hypertension. Because blood pressure is determined by the product of cardiac output and vascular resistance, candidate genes may act through either pathway. Physiologic pathways which are known to influence these parameters include the renin-angiotensin-aldosterone system, which contributes to determination of both cardiac output and vascular resistance. In this pathway, angiotensinogen, a hormone produced in the liver, is cleaved by an enzyme called renin to angiotensin I, which then undergoes further cleavage by angiotensin I-converting enzyme (ACE) to produce the active hormone angiotensin II (All). All acts through specific AT1 receptors present on vascular and adrenal cells. Receptors present on vascular cells cause vasoconstriction of blood vessels. Receptors present on adrenal cells cause release of the hormone aldosterone by the adrenal gland. This hormone acts on the mineralocorticoid receptor to cause increase sodium reabsorption largely through a renal epithelial sodium location. Other candidate genes are those of peripheral and central adrenergic pathways, which have dominant effects on cardiac iontropy, heart rate and vascular resistance; a variety of renal ion channels and transporters, which determine net sodium absorption and hence intravascular volume; calcium channels and exchangers and nitric oxide pathways, whose activity influences vascular tone. Another candidate gene encodes atrial natriuretic factor precursor, which is cleaved to atrial natriuretic peptides, found in the heart atrium, an endocrine organ controlling blood pressure and organ volume.

[0006] For some of the above candidate genes, variant forms have been identified that occur with increased frequency in individuals with hypertension. For example, a number of the polymorphisms have been reported in the angiotensinogen gene (AGT). In one of these, an M/T substitution at position 235, the T allele occurs more frequently in individuals with hypertension suggesting that this polymorphic form is a cause of hypertension or in equilibrium dislinkage with another polymorphism that is a cause. Jeunmaitre et al., Am. J. Hum. Genet. 60, 1448-1460 (1997). Two other genes within the renin-angiotensin-aldosterone system also have variant forms correlated with specific forms of hypertension, that is, aldosterone synthase gene and the gene encoding the β -subunit of the epithelial sodium channel induced by the mineralocorticoid receptor. Lifton et al., Proc. Natl. Acad. Sci. USA 92, 8548-8551 (1995).

45 [0007] Despite these developments, only a minute proportion of the total repository of polymorphisms in candidate genes for hypertension has been identified, and the primary genetic determinants of hypertension remain unknown in most affected subjects, as does the nature of the interaction between different genetic determinants. The paucity of polymorphisms hitherto identified is due to the large amount of work required for their detection by conventional methods. For example, a conventional approach to identifying polymorphisms might be to sequence the same stretch of oligonucleotides in a population of individuals by dideoxy sequencing. In this type of approach, the amount of work increases in proportion to both the length of sequence and the number of individuals in a population and becomes impractical for large stretches of DNA or large numbers of persons.

SUMMARY OF THE INVENTION

55 [0008] The invention provides nucleic acids of between 10 and 100 bases comprising at least 10 contiguous nucleotides including a polymorphic site from a sequence shown in Table 1, column 8 or the complement thereof. The nucleic acids can be DNA or RNA. Some nucleic acids are between 10 and 50 bases and some are between 20 and 50 bases. The base occupying the polymorphic site in such nucleic acids can be either a reference base shown in Table 1, column

3 or an alternative base shown in Table 1, column 5. In the some nucleic acids, the polymorphic site is occupied by a base that correlates with hypertension or susceptibility thereto. Some nucleic acids contain a polymorphic site having two polymorphic forms giving rise two different amino acids specified by the two codons in which the polymorphic site occurs in the two polymorphic forms.

[0009] The invention further provides allele-specific oligonucleotides that hybridize to a nucleic acid segment shown in Table 1, column 8 or its complement, including the polymorphic site. Such oligonucleotides are useful as probes or primers.

[0010] The invention further provides methods of analyzing a nucleic acid sequence. Such methods entail obtaining the nucleic acid from an individual; and determining a base occupying any one of the polymorphic sites shown in Table 1 or other polymorphic sites in equilibrium dislinkage therewith. Some methods determine a set of bases occupying a set of the polymorphic sites shown in Table 1. In some methods, the nucleic acid is obtained from a plurality of individuals, and a base occupying one of the polymorphic positions is determined in each of the individuals. Each individual is then tested for the presence of a disease phenotype, and correlating the presence of the disease phenotype with the base, particularly hypertension.

[0011] In another aspect, the invention provides nucleic acids comprising an isolated nucleic acid sequence of Table 1, column 8 or the complement thereof, wherein the polymorphic site within the sequence or its complement is occupied by a base other than the reference base show in Table 1, column 3. Such nucleic acids are useful, for example, in expression of variant proteins or production of transgenic animals.

[0012] The invention further provides methods of diagnosing a phenotype. Such methods entail determining which polymorphic form(s) are present in a sample from a subject at one or more polymorphic sites shown in Table 1, and diagnosing the presence of a phenotype correlated with the form(s) in the subject.

[0013] The invention also provides methods of screening polymorphic sites linked to polymorphic sites shown in Table 1 for suitability for diagnosing a phenotype. Such methods entail identifying a polymorphic site linked to a polymorphic site shown in Table 1, wherein a polymorphic form of the polymorphic site shown in Table 1 has been correlated with a phenotype. One then determines haplotypes in a population of individuals to indicate whether the linked polymorphic site has a polymorphic form in equilibrium dislinkage with the polymorphic form correlated with the phenotype.

BRIEF DESCRIPTION OF THE DRAWINGS

[0014] Figs. 1A and 1B depict computer systems suitable for storing and transmitting information relating to the polymorphisms of the invention.

DEFINITIONS

[0015] A nucleic acid can be DNA or RNA, and single- or double-stranded. Oligonucleotides can be naturally occurring or synthetic, but are typically prepared by synthetic means. Preferred nucleic acids of the invention include segments of DNA, or their complements including any one of the polymorphic sites shown in Table 1. The segments are usually between 5 and 100 contiguous bases, and often range from 5, 10, 12, 15, 20, or 25 nucleotides to 10, 15, 30, 25, 20, 50 or 100 nucleotides. Nucleic acids between 5-10, 5-20, 10-20, 12-30, 15-30, 10-50, 20-50 or 20-100 bases are common. The polymorphic site can occur within any position of the segment. The segments can be from any of the allelic forms of DNA shown in Table 1. For brevity in Table 1, the symbol T is used to represent both thymidine in DNA and uracil in RNA. Thus, in RNA oligonucleotides, the symbol T should be construed to indicate a uracil residue.

[0016] Hybridization probes are capable of binding in a base-specific manner to a complementary strand of nucleic acid. Such probes include nucleic acids, peptide nucleic acids, as described in Nielsen et al., Science 254, 1497-1500 (1991).

[0017] The term primer refers to a single-stranded oligonucleotide capable of acting as a point of initiation of template-directed DNA synthesis under appropriate conditions (i.e., in the presence of four different nucleoside triphosphates and an agent for polymerization, such as, DNA or RNA polymerase or reverse transcriptase) in an appropriate buffer and at a suitable temperature. The appropriate length of a primer depends on the intended use of the primer but typically ranges from 15 to 30 nucleotides. Short primer molecules generally require cooler temperatures to form sufficiently stable hybrid complexes with the template. A primer need not reflect the exact sequence of the template but must be sufficiently complementary to hybridize with a template. The term primer site refers to the area of the target DNA to which a primer hybridizes. The term primer pair means a set of primers including a 5' upstream primer that hybridizes with the 5' end of the DNA sequence to be amplified and a 3', downstream primer that hybridizes with the complement of the 3' end of the sequence to be amplified.

[0018] Linkage describes the tendency of genes, alleles, loci or genetic markers to be inherited together as a result of their location on the same chromosome, and can be measured by percent recombination between the two genes, alleles, loci or genetic markers. Loci occurring within 50 centimorgan of each other are linked. Some linked markers

occur within the same gene or gene cluster.

[0019] Polymorphism refers to the occurrence of two or more genetically determined alternative sequences or alleles in a population. A polymorphic marker or site is the locus at which divergence occurs. Preferred markers have at least two alleles, each occurring at frequency of greater than 1%, and more preferably greater than 10% or 20% of a selected population. A polymorphic locus may be as small as one base pair. Polymorphic markers include restriction fragment length polymorphisms, variable number of tandem repeats (VNTR's), hypervariable regions, minisatellites, dinucleotide repeats, trinucleotide repeats, tetranucleotide repeats, simple sequence repeats, and insertion elements such as Alu. The first identified allelic form is arbitrarily designated as the reference form and other allelic forms are designated as alternative or variant alleles. The allelic form occurring most frequently in a selected population is sometimes referred to as the wildtype form. Diploid organisms may be homozygous or heterozygous for allelic forms. A diallelic polymorphism has two forms. A triallelic polymorphism has three forms.

[0020] A single nucleotide polymorphism occurs at a polymorphic site occupied by a single nucleotide, which is the site of variation between allelic sequences. The site is usually preceded by and followed by highly conserved sequences of the allele (e.g., sequences that vary in less than 1/100 or 1/1000 members of the populations).

[0021] A single nucleotide polymorphism usually arises due to substitution of one nucleotide for another at the polymorphic site. A transition is the replacement of one purine by another purine or one pyrimidine by another pyrimidine. A transversion is the replacement of a purine by a pyrimidine or vice versa. Single nucleotide polymorphisms can also arise from a deletion of a nucleotide or an insertion of a nucleotide relative to a reference allele.

[0022] A set of polymorphisms means at least 2, and sometimes 5, 10, 20, 50 or more of the polymorphisms shown in Table 1.

[0023] Hybridizations are usually performed under stringent conditions, for example, at a salt concentration of no more than 1 M and a temperature of at least 25°C. For example, conditions of 5X SSPE (750 mM NaCl, 50 mM Na Phosphate, 5 mM EDTA, pH 7.4) and a temperature of 25-30°C are suitable for allele-specific probe hybridizations.

[0024] An isolated nucleic acid means an object species invention that is the predominant species present (i.e., on a molar basis it is more abundant than any other individual species in the composition). Preferably, an isolated nucleic acid comprises at least about 50, 80 or 90 percent (on a molar basis) of all macromolecular species present. Most preferably, the object species is purified to essential homogeneity (contaminant species cannot be detected in the composition by conventional detection methods).

[0025] Linkage disequilibrium or allelic association means the preferential association of a particular allele or genetic marker with a specific allele, or genetic marker at a nearby chromosomal location more frequently than expected by chance for any particular allele frequency in the population. For example, if locus X has alleles a and b, which occur equally frequently, and linked locus Y has alleles c and d, which occur equally frequently, one would expect the haplotype ac to occur with a frequency of 0.25 in a population of individuals. If ac occurs more frequently, then alleles a and c are in linkage disequilibrium. Linkage disequilibrium may result from natural selection of certain combination of alleles or because an allele has been introduced into a population too recently to have reached equilibrium with linked alleles.

[0026] A marker in linkage disequilibrium can be particularly useful in detecting susceptibility to disease (or other phenotype) notwithstanding that the marker does not cause the disease. For example, a marker (X) that is not itself a causative element of a disease, but which is in linkage disequilibrium with a gene (including regulatory sequences) (Y) that is a causative element of a phenotype, can be used detected to indicate susceptibility to the disease in circumstances in which the gene Y may not have been identified or may not be readily detectable. Younger alleles (i.e., those arising from mutation relatively late in evolution) are expected to have a larger genomic sequencing in linkage disequilibrium. The age of an allele can be determined from whether the allele is shared between ethnic human groups and/or between humans and related species.

DETAILED DESCRIPTION

[0027] The invention provides a substantial collection of novel polymorphisms in several genes encoding products known or suspected to have roles in biochemical pathways relating to blood pressure. Detection of polymorphisms in such genes is useful in designing and performing diagnostic assays for hypertension. Analysis of polymorphisms is also useful in designing prophylactic and therapeutic regimes customized to underlying abnormalities. As with other human polymorphisms, the polymorphisms of the invention also have more general applications, such as forensics, paternity testing, linkage analysis and positional cloning.

I. Novel Polymorphisms of the Invention.

[0028] The invention provides polymorphic sites in 75 candidate genes, known or suspected to have roles in hypertension. A gene was designated a candidate based on known or suggested involvement in blood pressure homeostasis and/or hypertension in one of the following biochemical pathways: renin-angiotensin, neural, or hormonal pathways

regulating blood pressure; regulation of vascular constriction, growth, and repair; ion and other small molecule transportation pathways in the kidney; and, regulation of glucose metabolism. Experimental evidence supporting selection of candidate genes included blood pressure physiology, animal models with altered blood pressure (including transgenic and knockout mouse or rat animal models), and human genetic linkage and association studies.

[0029] To maximize the chances of identifying informative single nucleotide polymorphisms (SNPs), DNA samples from 40 Africans and 35 US individuals of Northern European descent were screened to include both a range of human genetic diversity and hypertension phenotype diversity. Human genetic diversity is greater within African, as compared to European, Asian or American, populations (*The History and Geography of Human Genes* (Cavalli-Sforza et al., Eds., Princeton University Press, Princeton, NJ, 1994)). There are also significant differences in the prevalence and phenotype of hypertension between Africans (or US Blacks) and Northern Europeans (or US Whites). Hypertension has a greater prevalence, an earlier onset and a higher frequency of salt-sensitive cases in populations of African descent. The individuals sampled were selected from the top and bottom 2.5th percentile of a normalized blood pressure distribution. Regression analysis was performed within each community sample, of systolic, diastolic and mean arterial blood pressure against age and sex, and calculated the ranked frequency distribution of residuals. Equal numbers of individuals were selected from both ends of this latter distribution to maximize potential genetic differences if the genes screened for SNPs.

[0030] 874 SNPs in 75 individuals were identified at a frequency of one SNP per 217 bases. 387 SNPs were in coding sequences, 150 in introns, and 337 in 5' and 3' UTRs. Of coding sequence changes, 178 and 209 SNPs led to synonymous and nonsynonymous substitutions in the translated protein. On average, 12 SNPs were identified per gene, with the number ranging from zero (HSD11) to 54 (PGIS), with ten genes harboring 20 or more SNPs.

[0031] A large collection of polymorphisms of the invention are listed in Table 1. The first column of the Table 1 lists the gene and exon in which a given polymorphism occurs. For example, ACEEX13 means that a polymorphism occurs in exon 13 of angiotensin I-converting enzyme. AGTEX2 means that a polymorphism occurs in exon 2 of the angiotensinogen gene. The full names of the 75 genes shown in Table 1 are shown in Table 3. Sequences of each of the genes are available at <http://www.ncbi.nlm.nih.gov/Entrez/nucleotide.html>. The second column of Table 1 shows the position of a polymorphism. Numbering of nucleotides follows that of previously published reference sequences with nucleotides in sequence tags shown in column 8 being assigned the same number as the corresponding nucleotide in a reference sequence when the two are maximally aligned. In general, nucleotides in exons are numbered consecutively from the first base of the exon. Column 3 shows the base occupying the polymorphic position in a previously published sequence (arbitrarily designated a reference sequence). Column 4 of Table 1 shows the population frequency of the reference allele. For example at position 138 of exon 13 of ACE, a C nucleotide occurs in 63% of the population. Column 5 of the table shows a nucleotide occupying a polymorphic position that differs from previously published sequences. An allele containing such a nucleotide is designated an alternative allele. Column 6 of the Table shows the population frequency of the alternative allele. Column 7 of the Table shows the population frequency of heterozygosity at a polymorphic position. For example, for the polymorphic position at position 138 of exon 13 of the ACE gene, 37% of the human population are heterozygous. A high frequency of heterozygosity is advantageous in many applications of polymorphisms. The eighth column of the table shows a polymorphic position and about 15 nucleotides of flanking sequence on either side. The bases occupying the polymorphic position are indicated using IUPAC ambiguity nomenclature. For polymorphisms occurring in coding regions, columns 9 and 10 of the Table indicate the codons of the reference and alternate alleles including the polymorphic site. These columns are left blank for polymorphisms occurring in noncoding regions. Column 11 indicates whether the change between reference and alternate alleles is synonymous (i.e., no amino acid substitution due to polymorphic variation), nonsynonymous (i.e., polymorphic variation causes amino acid substitution). If the polymorphic site does not occur in a coding region, column 11 characterizes the polymorphic site as "other." For polymorphic sites occurring in noncoding regions column 12 indicates the type of region in which the site occurs (e.g., 5' UTR, intron). For polymorphic sites occurring in coding regions, column 12 indicates the amino acid encoded by the codon of the reference allele in which the polymorphic site occurs. Column 13 indicates the amino acid encoded by the codon of the alternative allele in which the polymorphic site occurs.

[0032] The polymorphisms shown in Tables 1 were identified by resequencing of target sequences from unrelated individuals of diverse ethnic and geographic backgrounds by hybridization to probes immobilized to microfabricated arrays. About 190 kb of genomic sequence from 75 candidate genes in 75 humans (150 alleles) or about 28 MB total was analyzed. The sequence included 87 kb coding DNA, 25 kb intron and 77 kb of 5' and 3' UTR sequences. Multiple target sequences from an individual were amplified from human genomic DNA using primers complementary to published sequences. The amplified target sequences were fluorescently labelled during or after PCR.

[0033] Polymorphisms were identified by hybridization of amplified DNA to arrays of oligonucleotide probes. Each genomic region was amplified by the polymerase chain reaction (PCR) in multiple segments, ranging from 80 bp to 14 kb, by both conventional and long PCR protocols. 205 distinct PCR products, averaging 3 kb, representing all 75 genes were pooled for each individual for each chip design.

[0034] The strategy and principles for design and use of arrays of oligonucleotide probes are generally described in

WO 95/11995. The strategy provides arrays of probes for analysis of target sequences showing a high degree of sequence identity to the published sequences described above. A typical probe array used in this analysis has two groups of four sets of probes that respectively tile both strands of a reference sequence. A first probe set comprises a plurality of probes exhibiting perfect complementarity with one of the reference sequences. Each probe in the first probe set has an interrogation position that corresponds to a nucleotide in the reference sequence. That is, the interrogation position is aligned with the corresponding nucleotide in the reference sequence, when the probe and reference sequence are aligned to maximize complementarity between the two. For each probe in the first set, there are three corresponding probes from three additional probe sets. Thus, there are four probes corresponding to each nucleotide in the reference sequence. The probes from the three additional probe sets are identical to the corresponding probe from the first probe set except at the interrogation position, which occurs in the same position in each of the four corresponding probes from the four probe sets, and is occupied by a different nucleotide in the four probe sets. Arrays tiled for multiple different reference sequences were included on the same substrate.

[0035] The labelled target sequences were hybridized with a substrate bearing immobilized arrays of probes. The amount of label bound to probes was measured. Analysis of the pattern of label revealed the nature and position of differences between the target and reference sequence. For example, comparison of the intensities of four corresponding probes reveals the identity of a corresponding nucleotide in the target sequences aligned with the interrogation position of the probes. The corresponding nucleotide is the complement of the nucleotide occupying the interrogation position of the probe showing the highest intensity (see WO 95/11995). The existence of a polymorphism is also manifested by differences in normalized hybridization intensities of probes flanking the polymorphism when the probes hybridized to corresponding targets from different individuals. For example, relative loss of hybridization intensity in a "footprint" of probes flanking a polymorphism signals a difference between the target and reference (i.e., a polymorphism) (see EP 717,113, incorporated by reference in its entirety for all purposes). Additionally, hybridization intensities for corresponding targets from different individuals can be classified into groups or clusters suggested by the data, not defined a priori, such that isolates in a give cluster tend to be similar and isolates in different clusters tend to be dissimilar. See WO 97/29212, filed February 7, 1997 (incorporated by reference in its entirety for all purposes). Hybridizations to samples from different individuals were performed separately.

II. Analysis of Polymorphisms

A. Preparation of Samples

[0036] Polymorphisms are detected in a target nucleic acid from an individual being analyzed. For assay of genomic DNA, virtually any biological sample (other than pure red blood cells) is suitable. For example, convenient tissue samples include whole blood, semen, saliva, tears, urine, fecal material, sweat, buccal, skin and hair. For assay of cDNA or mRNA, the tissue sample must be obtained from an organ in which the target nucleic acid is expressed.

[0037] Many of the methods described below require amplification of DNA from target samples. This can be accomplished by e.g., PCR. See generally PCR Technology: Principles and Applications for DNA Amplification (ed. H.A. Erlich, Freeman Press, NY, NY, 1992); PCR Protocols: A Guide to Methods and Applications (eds. Innis, et al., Academic Press, San Diego, CA, 1990); Mattila et al., Nucleic Acids Res. 19, 4967 (1991); Eckert et al., PCR Methods and Applications 1, 17 (1991); PCR (eds. McPherson et al., IRL Press, Oxford); and U.S. Patent 4,683,202 (each of which is incorporated by reference for all purposes).

[0038] Other suitable amplification methods include the ligase chain reaction (LCR) (see Wu and Wallace, Genomics 4, 560 (1989), Landegren et al., Science 241, 1077 (1988), transcription amplification (Kwoh et al., Proc. Natl. Acad. Sci. USA 86, 1173 (1989)), and self-sustained sequence replication (Guatelli et al., Proc. Nat. Acad. Sci. USA, 87, 1874 (1990)) and nucleic acid based sequence amplification (NASBA). The latter two amplification methods involve isothermal reactions based on isothermal transcription, which produce both single stranded RNA (ssRNA) and double stranded DNA (dsDNA) as the amplification products in a ratio of about 30 or 100 to 1, respectively.

B. Detection of Polymorphisms in Target DNA

[0039] The identity of bases occupying the polymorphic sites shown in Table 1 can be determined in an individual (e.g., a patient being analyzed) by several methods, which are described in turn.

1. Allele-Specific Probes

[0040] The design and use of allele-specific probes for analyzing polymorphisms is described by e.g., Saiki et al., Nature 324, 163-166 (1986); Dattagupta, EP 235,726, Saiki, WO 89/11548. Allele-specific probes can be designed that hybridize to a segment of target DNA from one individual but do not hybridize to the corresponding segment from

another individual due to the presence of different polymorphic forms in the respective segments from the two individuals. Hybridization conditions should be sufficiently stringent that there is a significant difference in hybridization intensity between alleles, and preferably an essentially binary response, whereby a probe hybridizes to only one of the alleles. Some probes are designed to hybridize to a segment of target DNA such that the polymorphic site aligns with a central position (e.g., in a 15 mer at the 7 position; in a 16 mer, at either the 8 or 9 position) of the probe. This design of probe achieves good discrimination in hybridization between different allelic forms.

[0041] Allele-specific probes are often used in pairs, one member of a pair showing a perfect match to a reference form of a target sequence and the other member showing a perfect match to a variant form. Several pairs of probes can then be immobilized on the same support for simultaneous analysis of multiple polymorphisms within the same target sequence.

2. Tiling Arrays

[0042] The polymorphisms can also be identified by hybridization to nucleic acid arrays, some example of which are described by WO 95/11995 (incorporated by reference in its entirety for all purposes). One form of such arrays is described in the Examples section in connection with de novo identification of polymorphisms. The same array or a different array can be used for analysis of characterized polymorphisms. WO 95/11995 also describes subarrays that are optimized for detection of a variant forms of a precharacterized polymorphism. Such a subarray contains probes designed to be complementary to a second reference sequence, which is an allelic variant of the first reference sequence. The second group of probes is designed by the same principles as described in the Examples except that the probes exhibit complementarity to the second reference sequence. The inclusion of a second group (or further groups) can be particularly useful for analyzing short subsequences of the primary reference sequence in which multiple mutations are expected to occur within a short distance commensurate with the length of the probes (i.e., two or more mutations within 9 to 21 bases).

3. Allele-Specific Primers

[0043] An allele-specific primer hybridizes to a site on target DNA overlapping a polymorphism and only primes amplification of an allelic form to which the primer exhibits perfect complementarity. See Gibbs, Nucleic Acid Res. 17, 2427-2448 (1989). This primer is used in conjunction with a second primer which hybridizes at a distal site. Amplification proceeds from the two primers leading to a detectable product signifying the particular allelic form is present. A control is usually performed with a second pair of primers, one of which shows a single base mismatch at the polymorphic site and the other of which exhibits perfect complementarity to a distal site. The single-base mismatch prevents amplification and no detectable product is formed. The method works best when the mismatch is included in the 3'-most position of the oligonucleotide aligned with the polymorphism because this position is most destabilizing to elongation from the primer. See, e.g., WO 93/22456.

4. Direct-Sequencing

[0044] The direct analysis of the sequence of polymorphisms of the present invention can be accomplished using either the dideoxy- chain termination method or the Maxam -Gilbert method (see Sambrook et al., Molecular Cloning, A Laboratory Manual (2nd Ed., CSHP, New York 1989); Zyskind et al., Recombinant DNA Laboratory Manual, (Acad. Press, 1988)).

5. Denaturing Gradient Gel Electrophoresis

[0045] Amplification products generated using the polymerase chain reaction can be analyzed by the use of denaturing gradient gel electrophoresis. Different alleles can be identified based on the different sequence-dependent melting properties and electrophoretic migration of DNA in solution. Erlich, ed., PCR Technology, Principles and Applications for DNA Amplification, (W.H. Freeman and Co, New York, 1992), Chapter 7.

6. Single-Strand Conformation Polymorphism Analysis

[0046] Alleles of target sequences can be differentiated using single-strand conformation polymorphism analysis, which identifies base differences by alteration in electrophoretic migration of single stranded PCR products, as described in Orita et al., Proc. Nat. Acad. Sci. 86, 2766-2770 (1989). Amplified PCR products can be generated as described above, and heated or otherwise denatured, to form single stranded amplification products. Single-stranded nucleic acids may refold or form secondary structures which are partially dependent on the base sequence. The different

electrophoretic mobilities of single-stranded amplification products can be related to base-sequence difference between alleles of target sequences.

III. Methods of Use

[0047] After determining polymorphic form(s) present in an individual at one or more polymorphic sites, this information can be used in a number of methods

A. Association Studies with Hypertension

[0048] The polymorphisms of the invention may contribute to the phenotype of an organism in different ways. Some polymorphisms occur within a protein coding sequence and contribute to phenotype by affecting protein structure. The effect may be neutral, beneficial or detrimental, or both beneficial and detrimental, depending on the circumstances. By analogy, a heterozygous sickle cell mutation confers resistance to malaria, but a homozygous sickle cell mutation is usually lethal. Other polymorphisms occur in noncoding regions but may exert phenotypic effects indirectly via influence on replication, transcription, and translation. A single polymorphism may affect more than one phenotypic trait. Likewise, a single phenotypic trait may be affected by polymorphisms in different genes. Further, some polymorphisms predispose an individual to a distinct mutation that is causally related to a certain phenotype.

[0049] The polymorphism shown in Table 1 are analyzed for a correlation with hypertension, the metabolic processes that lead to hypertension, and response to drugs used to treat hypertension. For purposes of these studies, hypertension can be defined as a dichotomous trait (e.g., diastolic blood pressure greater than 90 mm Hg), as a continuous scale of increasing severity based on blood pressure values, or as several intermediate phenotypes. Because it is likely that the causation of hypertension in the population is heterogenous, use of intermediate phenotypes can increase the strength of correlations identified. Some useful subtypes for association studies are mendelian forms of human hypertension, forms characterized by increased erythrocyte sodium-lithium countertransport, forms characterized by altered urinary kallikrein levels, and forms characterized by sensitivity of blood pressure to increases or decreases in sodium intake.

[0050] Correlation is performed for a population of individuals who have been tested for the presence or absence of hypertension or an intermediate phenotype and for one or polymorphic markers. To perform such analysis, the presence or absence of a set of polymorphic forms (i.e. a polymorphic set) is determined for a set of the individuals, some of whom exhibit a particular trait, and some of which exhibit lack of the trait. The alleles of each polymorphism of the set are then reviewed to determine whether the presence or absence of a particular allele is associated with the trait of interest. Correlation can be performed by standard statistical methods such as a K-squared test and statistically significant correlations between polymorphic form(s) and phenotypic characteristics are noted. For example, it might be found that the presence of allele A1 at polymorphism A correlates with hypertension as a dichotomous trait. As a further example, it might be found that the combined presence of allele A1 at polymorphism A and allele B1 at polymorphism B correlates with increased erythrocyte sodium lithium counter transport, an intermediate phenotype in development of hypertension.

B. Diagnosis of Hypertension

[0051] Polymorphic forms that correlate with hypertension or intermediate phenotypes are useful in diagnosing hypertension or susceptibility thereto. Combined detection of several such polymorphic forms (for example, 2, 5, 10 or 20 of the polymorphisms listed in Table 1) typically increases the probability of an accurate diagnosis. For example, the presence of a single polymorphic form known to correlate with hypertension might indicate a probability of 20% that an individual has or is susceptible to hypertension, whereas detection of five polymorphic forms, each of which correlates with hypertension, might indicate a probability of 80% that an individual has or is susceptible to hypertension. Analysis of the polymorphisms of the invention can be combined with that of other polymorphisms or other risk factors of hypertension, such as family history or obesity, as well as measurements of blood pressure.

[0052] Patients diagnosed with hypertension can be treated with conventional therapies and/or can be counselled to undertake remedial life style changes, such as a low fat, low salt diet or more exercise. Conventional therapies include diuretics (e.g., thiazides), which lower blood pressure by depleting the body of sodium and reducing blood volume; sympathoplegic agents (e.g., methyl dopa and clonidine), which lower blood pressure by reducing peripheral vascular resistance, inhibiting cardiac function and increasing venous pooling in capacitance vessels; direct vasodilators (e.g., hydralazine, minoxidil, diazoxide and sodium nitroprusside), which reduce pressure by relaxing vascular smooth muscle; agents that block production or action of angiotensin (e.g., captopril, enalapril and lisinopril), and thereby reduce peripheral vascular resistance; and adrenergic neuron blocking agents (e.g., guanethidine, reserpine, propranolol) which prevent release of norepinephrine. See, e.g., Basic and Clinical Pharmacology (Ed. Katzung, Ap-

pleton & Lange, CT, 1989).

C. Drug Screening

[0053] The polymorphism(s) showing the strongest correlation with hypertension within a given gene are likely to have a causative role in hypertension. Such a role can be confirmed by producing a transgenic animal expressing a human gene bearing such a polymorphism and determining whether the animal develops hypertension. Polymorphisms in coding regions that result in amino acid changes usually cause hypertension by decreasing, increasing or otherwise altering the activity of the protein encoded by the gene in which the polymorphism occurs. Polymorphisms in coding regions that introduce stop codons usually cause hypertension by reducing (heterozygote) or eliminating (homozygote) functional protein produced by the gene. Occasionally, stop codons result in production of a truncated peptide with aberrant activities relative to the full-length protein. Polymorphisms in regulatory regions typically cause hypertension by causing increased or decreased expression of the protein encoded by the gene in which the polymorphism occurs. Polymorphisms in intronic sequences can cause hypertension either through the same mechanism as polymorphisms in regulatory sequences or by causing altered spliced patterns resulting in an altered protein. For example, alternative splice patterns have been reported for the human angiotensin II receptor gene (Curnow et al., *Molecular Endocrinology* 9, 1250-1262 (1995)).

[0054] The precise role of polymorphisms in hypertension can be elucidated by several means. Alterations in expression levels of a protein (e.g., sodium-calcium ion channel) can be determined by measuring protein levels in samples groups of persons characterized as having or not having hypertension (or intermediate phenotypes). Alterations in enzyme activity (e.g., renin), can similarly be detected by assaying for enzyme activity in samples from the above groups of persons. Alterations in receptor transducing activity (e.g., angiotensin II receptor, β -3-adrenergic receptor or bradykinin receptor B2) can be detected by comparing receptor ligand binding, either in vitro or in a cellular expression system.

[0055] Having identified certain polymorphisms as having causative roles in hypertension, and having elucidated at least in general terms whether such polymorphisms increase or decrease the activity or expression level of associated proteins, customized therapies can be devised for classes of patients with different genetic subtypes of hypertension. For example, if a polymorphism in a given protein causes hypertension by increasing the expression level or activity of the protein, hypertension associated with the polymorphism can be treated by administering an antagonist of the protein. If a polymorphism in a given protein causes hypertension by decreasing the expression level or activity of a protein, the form of hypertension associated with the polymorphism can be treated by administering the protein itself, a nucleic acid encoding the protein that can be expressed in a patient, or an analog or agonist of the protein.

[0056] Agonists, antagonists can be obtained by producing and screening large combinatorial libraries. Combinatorial libraries can be produced for many types of compound that can be synthesized in a step by step fashion. Such compounds include polypeptides, beta-turn mimetics, polysaccharides, phospholipids, hormones, prostaglandins, steroids, aromatic compounds, heterocyclic compounds, benzodiazepines, oligomeric N-substituted glycines and oligocarbamates. Large combinatorial libraries of the compounds can be constructed by the encoded synthetic libraries (ESL) method described in Affymax, WO 95/12608, Affymax, WO 93/06121, Columbia University, WO 94/08051, Pharmaco-copeia, WO 95/35503 and Scripps, WO 95/30642 (each of which is incorporated by reference for all purposes). Peptide libraries can also be generated by phage display methods. See, e.g., Devlin, WO 91/18980. The libraries of compounds can be initially screened for specific binding to the protein for which agonists or antagonists are to be identified, or to its natural binding partner. Preferred agents bind with a $K_d < \mu M$. For example, for receptor ligand combinations, the assay can be performed using cloned receptor immobilized to a support such as a microtiter well and binding of compounds can be measured in competition with ligand to the receptor. Agonist or antagonist activity can then be assayed using a cellular reporter system or a transgenic animal model.

[0057] The polymorphisms of the invention are also useful for conducting clinical trials of drug candidates for hypertension. Such trials are performed on treated or control populations having similar or identical polymorphic profiles at a defined collection of polymorphic sites. Use of genetically matched populations eliminates or reduces variation in treatment outcome due to genetic factors, leading to a more accurate assessment of the efficacy of a potential drug.

D. Other Diseases

[0058] The polymorphisms in Table 1 can also be tested for association with other disease that have known but hitherto unmapped genetic components (e.g., agammaglobulinemia, diabetes insipidus, Lesch-Nyhan syndrome, muscular dystrophy, Wiskott-Aldrich syndrome, Fabry's disease, familial hypercholesterolemia, polycystic kidney disease, hereditary spherocytosis, von Willebrand's disease, tuberous sclerosis, hereditary hemorrhagica telangiectasia, familial colonic polyposis, Ehlers-Danlos syndrome, osteogenesis imperfecta, and acute intermittent porphyria). Phenotypic traits also include symptoms of, or susceptibility to, multifactorial diseases of which a component is or may be genetic,

such as autoimmune diseases, inflammation, cancer, diseases of the nervous system, and infection by pathogenic microorganisms. Some examples of autoimmune diseases include rheumatoid arthritis, multiple sclerosis, diabetes (insulin-dependent and non-independent), systemic lupus erythematosus and Graves disease. Some examples of cancers include cancers of the bladder, brain, breast, colon, esophagus, kidney, leukemia, liver, lung, oral cavity, ovary, pancreas, prostate, skin, stomach and uterus. Phenotypic traits also include characteristics such as longevity, appearance (e.g., baldness, obesity), strength, speed, endurance, fertility, and susceptibility or receptivity to particular drugs or therapeutic treatments.

[0059] Such correlations can be exploited in several ways. In the case of a strong correlation between a set of one or more polymorphic forms and a disease for which treatment is available, detection of the polymorphic form set in a human or animal patient may justify immediate administration of treatment, or at least the institution of regular monitoring of the patient. Detection of a polymorphic form correlated with serious disease in a couple contemplating a family may also be valuable to the couple in their reproductive decisions. For example, the female partner might elect to undergo in vitro fertilization to avoid the possibility of transmitting such a polymorphism from her husband to her offspring. In the case of a weaker, but still statistically significant correlation between a polymorphic set and human disease, immediate therapeutic intervention or monitoring may not be justified. Nevertheless, the patient can be motivated to begin simple life-style changes (e.g., diet, exercise) that can be accomplished at little cost to the patient but confer potential benefits in reducing the risk of conditions to which the patient may have increased susceptibility by virtue of variant alleles. Identification of a polymorphic set in a patient correlated with enhanced receptiveness to one of several treatment regimes for a disease indicates that this treatment regime should be followed.

E. Forensics

[0060] Determination of which polymorphic forms occupy a set of polymorphic sites in an individual identifies a set of polymorphic forms that distinguishes the individual. See generally National Research Council, The Evaluation of Forensic DNA Evidence (Eds. Pollard et al., National Academy Press, DC, 1996). The more sites that are analyzed the lower the probability that the set of polymorphic forms in one individual is the same as that in an unrelated individual. Preferably, if multiple sites are analyzed, the sites are unlinked. Thus, polymorphisms of the invention are often used in conjunction with polymorphisms in distal genes. Preferred polymorphisms for use in forensics are diallelic because the population frequencies of two polymorphic forms can usually be determined with greater accuracy than those of multiple polymorphic forms at multi-allelic loci.

[0061] The capacity to identify a distinguishing or unique set of forensic markers in an individual is useful for forensic analysis. For example, one can determine whether a blood sample from a suspect matches a blood or other tissue sample from a crime scene by determining whether the set of polymorphic forms occupying selected polymorphic sites is the same in the suspect and the sample. If the set of polymorphic markers does not match between a suspect and a sample, it can be concluded (barring experimental error) that the suspect was not the source of the sample. If the set of markers does match, one can conclude that the DNA from the suspect is consistent with that found at the crime scene. If frequencies of the polymorphic forms at the loci tested have been determined (e.g., by analysis of a suitable population of individuals), one can perform a statistical analysis to determine the probability that a match of suspect and crime scene sample would occur by chance.

[0062] $p(ID)$ is the probability that two random individuals have the same polymorphic or allelic form at a given polymorphic site. In diallelic loci, four genotypes are possible: AA, AB, BA, and BB. If alleles A and B occur in a haploid genome of the organism with frequencies x and y , the probability of each genotype in a diploid organism are (see WO 95/12607):

Homozygote: $p(AA) = x^2$

Homozygote: $p(BB) = y^2 = (1-x)^2$

Single Heterozygote: $p(AB) = p(BA) = xy = x(1-x)$

Both Heterozygotes: $p(AB+BA) = 2xy = 2x(1-x)$

[0063] The probability of identity at one locus (i.e., the probability that two individuals, picked at random from a population will have identical polymorphic forms at a given locus) is given by the equation:

$$p(ID) = (x^2)^2 + (2xy)^2 + (y^2)^2.$$

[0064] These calculations can be extended for any number of polymorphic forms at a given locus. For example, the probability of identity $p(ID)$ for a 3-allele system where the alleles have the frequencies in the population of x , y and z , respectively, is equal to the sum of the squares of the genotype frequencies:

$$p(ID) = x^4 + (2xy)^2 + (2yz)^2 + (2xz)^2 + z^4 + y^4$$

[0065] In a locus of n alleles, the appropriate binomial expansion is used to calculate $p(ID)$ and $p(exc)$.

5 [0066] The cumulative probability of identity ($\text{cum } p(ID)$) for each of multiple unlinked loci is determined by multiplying the probabilities provided by each locus.

$$\text{cum } p(ID) = p(ID1)p(ID2)p(ID3) \dots p(IDn)$$

10 [0067] The cumulative probability of non-identity for n loci (i.e. the probability that two random individuals will be different at 1 or more loci) is given by the equation:

$$\text{cum } p(\text{nonID}) = 1 - \text{cum } p(ID).$$

15 [0068] If several polymorphic loci are tested, the cumulative probability of non-identity for random individuals becomes very high (e.g., one billion to one). Such probabilities can be taken into account together with other evidence in determining the guilt or innocence of the suspect.

20 F. Paternity Testing

[0069] The object of paternity testing is usually to determine whether a male is the father of a child. In most cases, the mother of the child is known and thus, the mother's contribution to the child's genotype can be traced. Paternity testing investigates whether the part of the child's genotype not attributable to the mother is consistent with that of the putative father. Paternity testing can be performed by analyzing sets of polymorphisms in the putative father and the child.

25 [0070] If the set of polymorphisms in the child attributable to the father does not match the putative father, it can be concluded, barring experimental error, that the putative father is not the real father. If the set of polymorphisms in the child attributable to the father does match the set of polymorphisms of the putative father, a statistical calculation can be performed to determine the probability of coincidental match.

30 [0071] The probability of parentage exclusion (representing the probability that a random male will have a polymorphic form at a given polymorphic site that makes him incompatible as the father) is given by the equation (see WO 95/12607):

$$p(exc) = xy(1-xy)$$

where x and y are the population frequencies of alleles A and B of a diallelic polymorphic site.

40 [0072] (At a triallelic site $p(exc) = xy(1-xy) + yz(1-yz) + xz(1-xz) + 3xyz(1-xyz)$), where x , y and z are the respective population frequencies of alleles A, B and C).

[0073] The probability of non-exclusion is

$$p(\text{non-exc}) = 1 - p(exc)$$

45 [0074] The cumulative probability of non-exclusion (representing the value obtained when n loci are used) is thus:

$$\text{cum } p(\text{non-exc}) = p(\text{non-exc1})p(\text{non-exc2})p(\text{non-exc3}) \dots p(\text{non-exc}n)$$

50 [0075] The cumulative probability of exclusion for n loci (representing the probability that a random male will be excluded)

$$\text{cum } p(exc) = 1 - \text{cum } p(\text{non-exc}).$$

55 [0076] If several polymorphic loci are included in the analysis, the cumulative probability of exclusion of a random

male is very high. This probability can be taken into account in assessing the liability of a putative father whose polymorphic marker set matches the child's polymorphic marker set attributable to his/her father.

G. Genetic Mapping of Phenotypic Traits

[0077] The polymorphisms shown in table 1 can also be used to establish physical linkage between a genetic locus associated with a trait of interest and polymorphic markers that are not associated with the trait, but are in physical proximity with the genetic locus responsible for the trait and co-segregate with it. Such analysis is useful for mapping a genetic locus associated with a phenotypic trait to a chromosomal position, and thereby cloning gene(s) responsible for the trait. See Lander et al., *Proc. Natl. Acad. Sci. (USA)* 83, 7353-7357 (1986); Lander et al., *Proc. Natl. Acad. Sci. (USA)* 84, 2363-2367 (1987); Donis-Keller et al., *Cell* 51, 319-337 (1987); Lander et al., *Genetics* 121, 185-199 (1989). Genes localized by linkage can be cloned by a process known as directional cloning. See Wainwright, *Med. J. Australia* 159, 170-174 (1993); Collins, *Nature Genetics* 1, 3-6 (1992) (each of which is incorporated by reference in its entirety for all purposes).

[0078] Linkage studies are typically performed on members of a family. Available members of the family are characterized for the presence or absence of a phenotypic trait and for a set of polymorphic markers. The distribution of polymorphic markers in an informative meiosis is then analyzed to determine which polymorphic markers co-segregate with a phenotypic trait. See, e.g., Kerem et al., *Science* 245, 1073-1080 (1989); Monaco et al., *Nature* 316, 842 (1985); Yamoka et al., *Neurology* 40, 222-226 (1990); Rossiter et al., *FASEB Journal* 5, 21-27 (1991).

[0079] Linkage is analyzed by calculation of LOD (log of the odds) values. A lod value is the relative likelihood of obtaining observed segregation data for a marker and a genetic locus when the two are located at a recombination fraction θ , versus the situation in which the two are not linked, and thus segregating independently (Thompson & Thompson, *Genetics in Medicine* (5th ed, W.B. Saunders Company, Philadelphia, 1991); Strachan, "Mapping the human genome" in *The Human Genome* (BIOS Scientific Publishers Ltd, Oxford), Chapter 4). A series of likelihood ratios are calculated at various recombination fractions (θ), ranging from $\theta = 0.0$ (coincident loci) to $\theta = 0.50$ (unlinked). Thus, the likelihood at a given value of θ is: probability of data if loci linked at θ to probability of data if loci unlinked. The computed likelihoods are usually expressed as the log₁₀ of this ratio (i.e., a lod score). For example, a lod score of 3 indicates 1000:1 odds against an apparent observed linkage being a coincidence. The use of logarithms allows data collected from different families to be combined by simple addition. Computer programs are available for the calculation of lod scores for differing values of θ (e.g., LIPED, MLINK (Lathrop, *Proc. Nat. Acad. Sci. (USA)* 81, 3443-3446 (1984)). For any particular lod score, a recombination fraction may be determined from mathematical tables. See Smith et al., *Mathematical tables for research workers in human genetics* (Churchill, London, 1961); Smith, *Ann. Hum. Genet.* 32, 127-150 (1968). The value of θ at which the lod score is the highest is considered to be the best estimate of the recombination fraction. Positive lod score values suggest that the two loci are linked, whereas negative values suggest that linkage is less likely (at that value of θ) than the possibility that the two loci are unlinked. By convention, a combined lod score of +3 or greater (equivalent to greater than 1000:1 odds in favor of linkage) is considered definitive evidence that two loci are linked. Similarly, by convention, a negative lod score of -2 or less is taken as definitive evidence against linkage of the two loci being compared. Negative linkage data are useful in excluding a chromosome or a segment thereof from consideration. The search focuses on the remaining non-excluded chromosomal locations.

IV. Modified Polypeptides and Gene Sequences

[0080] The invention further provides variant forms of nucleic acids and corresponding proteins. The nucleic acids comprise one of the sequences described in Table 1, column 8, in which the polymorphic position is occupied by an alternative base for that position. Some nucleic acid encode full-length variant forms of proteins. Similarly, variant proteins have the prototypical amino acid sequences of encoded by nucleic acid sequence shown in Table 1, column 8, (read so as to be in-frame with the full-length coding sequence of which it is a component) except at an amino acid encoded by a codon including one of the polymorphic positions shown in the Table. That position is occupied by the amino acid coded by the corresponding codon in the alternative forms shown in the Table.

[0081] Variant genes can be expressed in an expression vector in which a variant gene is operably linked to a native or other promoter. Usually, the promoter is a eukaryotic promoter for expression in a mammalian cell. The transcription regulation sequences typically include a heterologous promoter and optionally an enhancer which is recognized by the host. The selection of an appropriate promoter, for example trp, lac, phage promoters, glycolytic enzyme promoters and tRNA promoters, depends on the host selected. Commercially available expression vectors can be used. Vectors can include host-recognized replication systems, amplifiable genes, selectable markers, host sequences useful for insertion into the host genome, and the like.

[0082] The means of introducing the expression construct into a host cell varies depending upon the particular construction and the target host. Suitable means include fusion, conjugation, transfection, transduction, electroporation

or injection, as described in Sambrook, supra. A wide variety of host cells can be employed for expression of the variant gene, both prokaryotic and eukaryotic. Suitable host cells include bacteria such as *E. coli*, yeast, filamentous fungi, insect cells, mammalian cells, typically immortalized, e.g., mouse, CHO, human and monkey cell lines and derivatives thereof. Preferred host cells are able to process the variant gene product to produce an appropriate mature polypeptide. Processing includes glycosylation, ubiquitination, disulfide bond formation, general post-translational modification, and the like.

[0083] The protein may be isolated by conventional means of protein biochemistry and purification to obtain a substantially pure product, i.e., 80, 95 or 99% free of cell component contaminants, as described in Jacoby, *Methods in Enzymology* Volume 104, Academic Press, New York (1984); Scopes, *Protein Purification, Principles and Practice*, 2nd Edition, Springer-Verlag, New York (1987); and Deutscher (ed), *Guide to Protein Purification, Methods in Enzymology*, Vol. 182 (1990). If the protein is secreted, it can be isolated from the supernatant in which the host cell is grown. If not secreted, the protein can be isolated from a lysate of the host cells.

[0084] The invention further provides transgenic nonhuman animals capable of expressing an exogenous variant gene and/or having one or both alleles of an endogenous variant gene inactivated. Expression of an exogenous variant gene is usually achieved by operably linking the gene to a promoter and optionally an enhancer, and microinjecting the construct into a zygote. See Hogan et al., "Manipulating the Mouse Embryo, A Laboratory Manual," Cold Spring Harbor Laboratory. Inactivation of endogenous variant genes can be achieved by forming a transgene in which a cloned variant gene is inactivated by insertion of a positive selection marker. See Capecchi, *Science* 244, 1288-1292 (1989). The transgene is then introduced into an embryonic stem cell, where it undergoes homologous recombination with an endogenous variant gene. Mice and other rodents are preferred animals. Such animals provide useful drug screening systems.

[0085] In addition to substantially full-length polypeptides expressed by variant genes, the present invention includes biologically active fragments of the polypeptides, or analogs thereof, including organic molecules which simulate the interactions of the peptides. Biologically active fragments include any portion of the full-length polypeptide which confers a biological function on the variant gene product, including ligand binding, and antibody binding. Ligand binding includes binding by nucleic acids, proteins or polypeptides, small biologically active molecules, or large cellular structures.

[0086] Polyclonal and/or monoclonal antibodies that specifically bind to variant gene products but not to corresponding prototypical gene products are also provided. Antibodies can be made by injecting mice or other animals with the variant gene product or synthetic peptide fragments thereof. Monoclonal antibodies are screened as are described, for example, in Harlow & Lane, *Antibodies, A Laboratory Manual*, Cold Spring Harbor Press, New York (1988); Goding, *Monoclonal antibodies, Principles and Practice* (2d ed.) Academic Press, New York (1986). Monoclonal antibodies are tested for specific immunoreactivity with a variant gene product and lack of immunoreactivity to the corresponding prototypical gene product. These antibodies are useful in diagnostic assays for detection of the variant form, or as an active ingredient in a pharmaceutical composition.

V. Kits

[0087] The invention further provides kits comprising at least one allele-specific oligonucleotide as described above. Often, the kits contain one or more pairs of allele-specific oligonucleotides hybridizing to different forms of a polymorphism. In some kits, the allele-specific oligonucleotides are provided immobilized to a substrate. For example, the same substrate can comprise allele-specific oligonucleotide probes for detecting at least 10, 100 or all of the polymorphisms shown in Table 1. Optional additional components of the kit include, for example, restriction enzymes, reverse-transcriptase or polymerase, the substrate nucleoside triphosphates, means used to label (for example, an avidin-enzyme conjugate and enzyme substrate and chromogen if the label is biotin), and the appropriate buffers for reverse transcription, PCR, or hybridization reactions. Usually, the kit also contains instructions for carrying out the methods.

VI. Computer Systems For Storing Polymorphism Data

[0088] Fig. 1A depicts a block diagram of a computer system 10 suitable for implementing the present invention. Computer system 10 includes a bus 12 which interconnects major subsystems such as a central processor 14, a system memory 16 (typically RAM), an input/output (I/O) controller 18, an external device such as a display screen 24 via a display adapter 26, serial ports 28 and 30, a keyboard 32, a fixed disk drive 34 via a storage interface 35 and a floppy disk drive 36 operative to receive a floppy disk 38, and a CD-ROM (or DVD-ROM) device 40 operative to receive a CD-ROM 42. Many other devices can be connected such as a user pointing device, e.g., a mouse 44 connected via serial port 28 and a network interface 46 connected via serial port 30.

[0089] Many other devices or subsystems (not shown) may be connected in a similar manner. Also, it is not necessary for all of the devices shown in Fig. 1A to be present to practice the present invention, as discussed below. The devices and subsystems may be interconnected in different ways from that shown in Fig. 1A. The operation of a computer

system such as that shown in Fig. 1A is well known. Databases storing polymorphism information according to the present invention can be stored, e.g., in system memory 16 or on storage media such as fixed disk 34, floppy disk 38, or CD-ROM 42. An application program to access such databases can be operably disposed in system memory 16 or sorted on storage media such as fixed disk 34, floppy disk 38, or CD-ROM 42.

5 [0090] Fig. 1B depicts the interconnection of computer system 10 to remote computers 48, 50, and 52. Fig. 1B depicts a network 54 interconnecting remote servers 48, 50, and 52. Network interface 46 provides the connection from client computer system 10 to network 54. Network 54 can be, e.g., the Internet. Protocols for exchanging data via the Internet and other networks are well known. Information identifying the polymorphisms described herein can be transmitted across network 54 embedded in signals capable of traversing the physical media employed by network 54.

10 [0091] Information identifying polymorphisms shown in Table 1 is represented in records, which optionally, are subdivided into fields. Each record stores information relating to a different polymorphism in Table 1. Collectively, the records can store information relating to all of the polymorphisms in Table 1, or any subset thereof, such as 5, 10, 50, or 100 polymorphisms from Table 1. In some databases, the information identifies a base occupying a polymorphic position and the location of the polymorphic position. The base can be represented as a single letter code (i.e., A, C, G or T/U) present in a polymorphic form other than that in the reference allele. Alternatively, the base occupying a polymorphic site can be represented in IUPAC ambiguity code as shown in Table 1. The location of a polymorphic site can be identified as its position within one of the sequences shown in Table 1. For example, in the first sequence shown in Table 1, the polymorphic site occupies the 15th base. The position can also be identified by reference to, for example, a chromosome, and distance from known markers within the chromosome. In other databases, information identifying a polymorphism contains sequences of 10-100 bases shown in Table 1 or the complements thereof including a polymorphic site. Preferably, such information records at least 10, 15, 20, or 30 contiguous bases of sequences including a polymorphic site.

15 [0092] From the foregoing, it is apparent that the invention includes a number of general uses that can be expressed concisely as follows. The invention provides for the use of any of the nucleic acid segments described above in the diagnosis or monitoring of diseases, particularly hypertension. The invention further provides for the use of any of the nucleic acid segments in the manufacture of a medicament for the treatment or prophylaxis of such diseases. The invention further provides for the use of any of the DNA segments as a pharmaceutical.

20 [0093] All publications and patent applications cited above are incorporated by reference in their entirety for all purposes to the same extent as if each individual publication or patent application were specifically and individually indicated to be so incorporated by reference. Although the present invention has been described in some detail by way of illustration and example for purposes of clarity and understanding, it will be apparent that certain changes and modifications may be practiced within the scope of the appended claims.

TABLE 1

Gene/Exon	Base Position	Ref Allele	Freq (P)	All Allele	Freq (Q)	Heteroz yosity (H)	Sequence Tag	Ref Codon	All Codon	Type of amino acid change	Ref amino acid	All amino acid
AADEX1	305	G	0.98	A	0.03	0.05	ACGGGGCGGAGCCRGAGCCGGAGCCGAC	GGG	TGG	Other	5' UTR	.
AADEX10	246	G	0.86	T	0.14	0.23	AAGCTTCGAGGAAGCCAGAAATGGAAGC	AAT	ATT	Nonsynonymous	Gly	Trp
AADEX12	43	A	0.94	T	0.06	0.11	GATCCGAGAGCAGAWTTTACAGGACATTA	TCT	TGT	Nonsynonymous	Asn	Ile
AADEX13	173	C	0.70	G	0.30	0.42	GAACGAGAAGGGCTGTGAAGGTGAGTGCT	.	.	Nonsynonymous	Ser	Cys
AADEX15	74	C	0.73	T	0.27	0.40	CCTAGTAAGTACCGYGCTGCTCCGCTCT	.	.	Other	Intron	.
AADEX16	1071	G	0.99	A	0.01	0.02	ATTCTGTCTATAGGAAAGGTATATCAGGA	.	.	Other	3' UTR	.
AADEX16	1321	C	0.98	T	0.02	0.04	GCCTGGGGCCCTTYGACATCACCGTCAT	.	.	Other	3' UTR	.
AADEX16	1328	A	0.91	G	0.09	0.17	GGCCCTCGACATCACCCTGCTATTGATGGA	.	.	Other	3' UTR	.
AADEX16	1478	A	0.89	G	0.11	0.19	CAGCTGACTAGGTRCAGGCAAGCTTG	.	.	Other	3' UTR	.
AADEX16	691	C	0.99	G	0.01	0.02	CAGCTTGGCTGCASGTACCCCTCCTGAG	.	.	Other	3' UTR	.
AADEX16	995	C	0.94	T	0.06	0.11	TATGCACTGCTGACYGACATCCCTCGAC	.	.	Other	3' UTR	.
AADEX2	31	A	0.98	G	0.03	0.05	TTTGATTCTGTAGGRACCTAGAAAGATTG	.	.	Other	5' UTR	.
AADEX7	96	T	0.98	A	0.02	0.04	TTGGAGAACTGGCTWATCATGACTACCAT	TAT	AAT	Nonsynonymous	Tyr	Asn
AADEX9	173	A	0.93	T	0.07	0.13	ATTGGTGAGCAGGAWTTTGAAGCCCTCAT	GAA	GAT	Nonsynonymous	Glu	Asp
ACEEX13	151	T	0.75	C	0.25	0.38	CCAGCCAGGAGGCAACCCCAACAGGTGACA	TCT	CCT	Nonsynonymous	Ser	Pro
ACEEX13	202	A	0.75	G	0.25	0.38	AGGCAACAACAGCAGCAGCAACACACC	AGC	GCC	Nonsynonymous	Ser	Gly
ACEEX15	144	G	0.80	A	0.20	0.32	CTAGAACGGGCAACGCTGCTGCCAGGA	CCG	GCA	Synonymous	Ala	Ala
ACEEX17	19	C	0.69	A	0.31	0.43	CTCAAGCATTCAAMCCCTTACAGATCT	.	.	Other	Intron	.
ACEEX18	130	C	0.95	G	0.05	0.09	CAGCCACTCTACCTSAACCTGCATGCCCTA	CTC	CTG	Synonymous	Leu	Leu
ACEEX21	150	T	0.98	C	0.03	0.05	CTTCCATGAGGCCAYTGGGACGTGCTAG	ATT	ACT	Nonsynonymous	Ile	Thr
ACEEX22	19	T	0.99	G	0.01	0.03	AGCATGACATCAACKTCTCTGATGAAGATG	TTT	GTT	Nonsynonymous	Phe	Val
ACEEX24	118	C	0.95	T	0.05	0.10	CAGTCCAAGGAGGCGYGGCAGCGCTGGG	GCC	GCT	Synonymous	Ala	Ala
ACEEX24	16	T	0.57	C	0.43	0.49	TGCTCCAGGTACTTGTGTCAGCTTCATCAT	TTT	TTC	Synonymous	Phe	Phe
ACEEX26	154	G	0.98	A	0.03	0.05	GGGCTCAGCCAGCAGCTCTTCAGCATCC	CGG	CAG	Nonsynonymous	Arg	Gln
ACEEX26	174	C	0.90	A	0.10	0.18	TCAGCATCCGCCACMCGAGCTTCACCGG	CGC	AGC	Nonsynonymous	Arg	Ser

55	ACEEX26	205	A	0.98	C	0.02	0.03	CTCCACGGGCCCCGTTTCGGCTCCGAGG	CAG	CCG	Nonsynonymous	Gln	Pro
	ACEEX26	224	G	0.94	A	0.06	0.11	GGCTCCGAGGTGGARCTGAGACACTCCTG	GAG	GAA	Synonymous	Glu	Glu
	ADJUEX10	81	G	0.98	T	0.03	0.05	CTCTGGAGCAGGAKAAGACACCGGCCCA	GAG	GAT	Nonsynonymous	Glu	Asp
	ADDBEX15	68	G	0.99	A	0.01	0.03	GCTCTGGTCCGCCCTGTGCGAGTTCTTC	CCG	CCA	Synonymous	Pro	Pro
	ADDBEX15	85	C	0.90	T	0.10	0.18	TGCGAGTTCTTCAGYGTTCGCCCTCCACAT	GCG	GTC	Nonsynonymous	Ala	Val
	ADDBEX17	147	C	0.98	A	0.02	0.04	GAGAAATCCTCAGMAAAGGCTGAGCCA	AGC	AGA	Nonsynonymous	Ser	Arg
	ADDBEX3	138	G	0.89	A	0.11	0.19	GCTTCTCAGAGACACACCCGAGTACATG	GAC	AAC	Nonsynonymous	Asp	Asn
	ADDBEX4	134	C	0.99	G	0.01	0.02	CATGGCCAGCACTSCCAGCCAGTCTTCC	TCC	TGC	Nonsynonymous	Ser	Cys
	ADDBEX8	173	A	0.96	G	0.04	0.07	CCACCTGCAAGTTTGTCTTAGCTCTCTGT	TTT	TTC	Other	Intron	
	ADDBEX9	69	T	0.99	C	0.01	0.02	GTAGAGAGGCAATYTTACAAAGATCTTCCA	TTT	TTC	Synonymous	Phe	Phe
	ADDB	2087	T	0.98	C	0.02	0.03	TGACATTGCACATCYAAATACCAATTTA	TTT	TTC	Other	3' UTR	
	ADORA2AEX1	429	C	0.97	T	0.03	0.07	GGTGCTACTGGCGGCGGCGGACATCGCAG	GCG	GTC	Nonsynonymous	Ala	Val
	ADORA2AEX2	1230	G	0.97	T	0.03	0.06	TGCAGAGCACTTCCRCAAAGATCATTCGCA	CGC	CAC	Other	3' UTR	
	ADORA2AEX2	596	G	0.97	A	0.03	0.06	CCGCCAGACCTTCCRCAAAGATCATTCGCA	CGC	CAC	Nonsynonymous	Arg	His
	ADORA2AEX2	741	C	0.92	T	0.08	0.15	GGAGTGTGGGCAAYGGCAGTGTCTCCCA	AAC	AAT	Synonymous	Asn	Asn
	ADRB3EX1	1020	C	0.96	T	0.04	0.07	GGCCCCGTGGGAYGTGCGTCTCGCCCG	ACG	ATG	Nonsynonymous	Thr	Met
	ADRB3EX1	1354	C	0.89	T	0.11	0.20	TGCGCCGCCGCCGCGCCGCCCTCTTCCC	CGC	CGT	Synonymous	Arg	Arg
	ADRB3EX1	1445	G	0.90	T	0.10	0.18	GGTAGGTAAACCGGKACAGGGACCGCGG	CGC	CGT	Other	Intron	
	ADRB3EX1	44	A	0.96	G	0.04	0.08	GCTACTCTCCCTCCCTCAGAGCGGTGGCAC	CGC	CGT	Other	5' UTR	
	ADRB3EX2	301	G	0.96	C	0.04	0.08	GTGGTAGTGTCCAGSTGCGTGGAGCAGC	CGC	CGT	Other	3' UTR	
	ADRB3EX2	408	C	0.80	T	0.20	0.32	TGGTTCATTCCTTCTTCTTCCACCCAAACCC	CGC	CGT	Other	3' UTR	
	ADROMEX1	1197	C	0.98	T	0.02	0.04	TGGAGCTCTGAGAYTTTCTCTCTCAAGT	CGC	CGT	Other	3' UTR	
	ADROMEX1	154	G	0.98	T	0.02	0.05	ATGTTACCTTCTCTTCTCTGACTCAAGGT	CGC	CGT	Other	3' UTR	
	ADROMEX1	723	C	0.97	T	0.03	0.06	GGCTCTCTGCTGTCTTCTCTGCGCAGGAGCT	CGC	CGT	Other	Promoter	
	ADROMEX1	981	G	0.99	A	0.01	0.03	GAGCAGGAGCGCGCTGCTGAGGAAAGA	CGC	CGT	Other	Promoter	
	ADROMEX2	101	A	0.96	C	0.04	0.07	TGCTCGCTTCTCTCTCTCTCTCTCTCTCTCT	CGC	CGT	Other	Promoter	
	ADROMEX3	81	C	0.95	G	0.05	0.09	CTGCGGATGCCAGSAGCTACCCACCCGG	CGC	CGT	Synonymous	Leu	Leu
	ADROMEX4	1033	T	0.95	C	0.05	0.09	ACCGAGTCTCTGTA YAA TCTA TTTACATA	CGC	CGT	Nonsynonymous	Ser	Arg
	ADROMEX4	1292	A	0.98	G	0.02	0.05	TGCTCTGGTGGCGARTCAGGCGTTCGCGG	CGC	CGT	Other	3' UTR	
	ADROMEX4	1389	T	0.97	C	0.03	0.06	UCGAGCTGGA CTCTCTCTCTCTCTCTCTCT	CGC	CGT	Other	3' UTR	
	ADROMEX4	388	G	0.98	C	0.02	0.05	CAAGCATCCCGCTCTCTCTCTCTCTCTCTCT	CGC	CGT	Other	3' UTR	

55	AGTEX2	181	C	0.99	T	0.01	0.03	AGAGTACCTGTGAGYAGCTGGCAAGGCC	CAG	TAG	Nonsynonymous	Gln	STOP
	AGTEX2	354	C	0.99	T	0.01	0.03	GTGGGATGCTGGCYAACCTCTTGGGCTT	GCC	GCT	Synonymous	Ala	Ala
	AGTEX2	755	T		G			GRACCTTACAGAAKCKGATGTGCTGCTG	CTG	CGG	Nonsynonymous	Leu	Arg
	AGTEX5	258	C	0.96	T	0.04	0.08	TGGCAAGCCCTCTGYCCCTGGCCCTTTGAG	.	.	Other	3' UTR	.
	AGTEX5	376	C	0.97	G	0.03	0.06	AGCTGAAAGCAGCGGTTCTCTTGGTC	.	.	Other	3' UTR	.
	AGTEX5	385	T	0.97	C	0.03	0.06	GCAGCCGTTTCYCTGGTCTAAGTGTC	.	.	Other	3' UTR	.
	AGTEX5	641	T	0.93	G	0.07	0.13	GCCTCGGTTTGTAATTAAGTGCTCTGAAT	.	.	Other	3' UTR	.
	AGTEXP1	101	G	0.99	C	0.01	0.03	CTGGCTGTGCTATTSTTGGTGTTTAACAG	.	.	Other	Promoter	.
	AGTEXP2	160	G	0.99	A	0.01	0.03	GGAACTTGGCCCRACCTCTGCAAACTT	.	.	Other	Promoter	.
	AGTEXP2	35	G	0.97	A	0.03	0.06	CCCTCTGCACCTCCRGCTGCACTGCTCT	.	.	Other	Promoter	.
	AGTEXP3	158	A	0.71	G	0.29	0.41	CTCTGACCCCGCCRGCGGAAGAAGCTGC	.	.	Other	Promoter	.
	AGTEXP3	173	C	0.96	T	0.04	0.08	GGGGAAGAAGCTGCTGCTCTCTGGGTAC	.	.	Other	Promoter	.
	ALDREDEX1	162	A	0.86	T	0.14	0.23	GGCCAAAGATGCCCTTCTTGGGTGGGT	ATC	TTC	Nonsynonymous	Ile	Phe
	ALDREDEX1	71	C	0.41	G	0.59	0.48	AAAGGTACGCCCGCGGCCAAGCCGCAC	.	.	Other	Promoter	.
	ALDREDEX10	150	T	0.91	G	0.09	0.16	TTGCAATGTAGTAKGGCTGTGTCACTC	.	.	Other	3' UTR	.
	ALDREDEX2	180	C	0.94	G	0.06	0.11	TGAAGCTGAGGAGSTCTTCATCGTGAGC	CTC	GTC	Nonsynonymous	Leu	Val
	ALDREDEX2	204	T	0.95	G	0.05	0.10	TCAGCAAGGTATCCKTCGCGGTGGGCT	.	.	Other	Intron	.
	ALDREDEX2	88	A	0.98	T	0.03	0.05	CGTCGGGTACGCCWCATCGACTGTGCCCC	CAC	CTC	Nonsynonymous	His	Leu
	ALDREDEX3	28	A	0.95	T	0.05	0.10	CCCTCUCCTGCTTTCCTGCTGCTGCTGCT	.	.	Other	Intron	.
	ALDREDEX4	101	G	0.98	A	0.03	0.05	AACATTCTGGACACRTGGCGGGTAAGACA	ACG	ACA	Synonymous	Thr	Thr
	ALDREDEX6	87	G	0.94	A	0.06	0.11	ACTGCCAGTCCAAAGCATCGTGCTGACC	GCC	AGC	Nonsynonymous	Gly	Ser
	ALDREDEX9	67	C	0.99	T	0.01	0.02	CCAGGATATGACCAVCTTACTCAGCTACA	ACC	ATC	Nonsynonymous	Thr	Ile
	ANPEX1	252	G	0.99	A	0.01	0.03	CCATGTACAATGCCRTGTCTCAACGCAGAC	GTG	ATG	Nonsynonymous	Val	Met
	ANPEX1	297	C	0.97	T	0.03	0.06	TAGGGCCAGGAAAGYGGGTGCAGTCTGGG	.	.	Other	Intron	.
	ANPEX3	106	G	0.97	T	0.03	0.06	TCCTGTCCCTGGGCTCTCTGCTGCATTT	.	.	Other	3' UTR	.
	ANPEX3	127	T	0.91	C	0.09	0.16	CTGCATTTGTGTCAVCTTGTTCGCATGGA	.	.	Other	3' UTR	.
	APOA1	101	C	0.76	T	0.24	0.36	GCCTTGGCCAGGCGYGGGCTCTGGGTAC	.	.	Other	Promoter	.
	APOA1	1016	A	0.76	C	0.24	0.36	CGTAAGTGGGACCCMGTCACAGCTCTGTC	.	.	Other	Intron	.
	APOA1	1162	G	0.94	C	0.06	0.12	AUGTGTCAACCCAGGSCCTCACCCCTGATAG	.	.	Other	Intron	.
	APOA1	1163	C	0.93	T	0.08	0.14	GGTGTACCCAGGGGTACCCCTGTATAGG	.	.	Other	Intron	.
	APOA1	1401	G	0.99	C	0.01	0.02	TGCAGCCCTACCTGSAGACCTTCCAGAAAG	GAC	CAC	Nonsynonymous	Asp	His

BNSDOCID: <EP_____09553B2A2_I_>

55	APOA4	2837	G	0.98	C	0.02	0.04	AGTTCCGACGCCGGSTGGAGCCCTACGGG	GTG	CTG	Nonsynonymous	Val	Leu
	APOA4	2926	G	0.81	T	0.19	0.30	CATCGGGGACGTGGAAGGCCACTTGAG	GTG	GTT	Synonymous	Val	Val
	APOA4	3058	T	0.16	G	0.84	0.26	CAGCAGGAACAGCAKACAGGAGCAGCACA	CAT	CAG	Nonsynonymous	Ile	Gln
	APOA4	350	G	0.88	A	0.12	0.21	GCCAGCAGGGCTCRAAGGCA TCAGTCCCG	.	.	Other	Promoter	.
	APOA4	637	G	0.96	C	0.04	0.08	TGGCGATAAGGAGASAGTTTAAATCTCTG	.	.	Other	Promoter	.
	APOA4	687	G	0.96	A	0.04	0.07	GTTCACACTGCAGCRCAGGTGAGCTCTCC	.	.	Other	Promoter	.
	APOC1EX1	1020	G	0.93	T	0.07	0.13	TTGTATTTTCAGTAKAGACAGGGTTTCAC	.	.	Other	Intron	.
	APOC1EX1	1044	G	0.95	A	0.05	0.10	TTCAACGTGGTCTCNA TCCTGTGACTTTG	.	.	Other	Intron	.
	APOC1EX1	1057	T	0.64	C	0.36	0.46	CGATCTCTGACTTGTGTGATCCGCCCTGCC	.	.	Other	Intron	.
	APOC1EX1	1111	C	0.89	T	0.11	0.20	CAGCGTGAGCCACACACAGGCTCCGGCAATTC	.	.	Other	Intron	.
	APOC1EX1	1376	G	0.57	T	0.43	0.49	GCACGCCCTGTAGKCCCAAGCTACTCCGG	.	.	Other	Intron	.
	APOC1EX1	1411	C	0.99	G	0.01	0.02	AGGCAGGAGATCASTTGAACCCCGGAGG	.	.	Other	Intron	.
	APOC1EX1	432	G	0.97	A	0.03	0.06	AGGCTCTCTGTCTCTCTCCCGGTCTCTGGT	TGG	TCA	Synonymous	Ser	Ser
	APOC1EX1	462	C	0.61	G	0.39	0.47	GTGGTCTCTCGATGTCCTTGGAAAGGTAA	ATC	ATG	Nonsynonymous	Ile	Met
	APOC1EX1	496	G	0.01	C	0.99	0.02	GGA TGGAGAA TGGGAGATTTGGAGATT	.	.	Other	Intron	.
	APOC1EX1	713	C	0.99	T	0.01	0.02	ACCTCTGGGATTTGGYTGCTCTCTCTCGAC	.	.	Other	Intron	.
	APOC2	1084	T	0.91	G	0.09	0.17	TCTGAGGACTCAAGKGCACAGATGGAGGG	.	.	Other	3' UTR	.
	APOC2	126	C	0.99	T	0.01	0.02	CAGGCTCTTGGACATYATGGGCACACGAC	.	.	Other	5' UTR	.
	APOC2	13	T	0.34	A	0.66	0.45	CTGGGACACCGAGCWCACACAGAGCAAGA	.	.	Other	Promoter	.
	APOC2	472	G	0.99	A	0.01	0.02	CCAGAACTCTGTACRAGAAGACATACCTG	GAG	AAG	Nonsynonymous	Glu	Lys
	APOC2	553	G	0.99	A	0.01	0.02	TGGCCCATACCCACRACCTGCCATCCAGGAC	.	.	Other	Intron	.
	APOC2	725	T	0.19	C	0.81	0.31	CCAGAGACTCCAGGVCCECCAGACCTCTCT	.	.	Other	Intron	.
	APOC2	804	A	0.97	T	0.03	0.06	TGTGCTTCTCCCCWGGGACTTTGTACAGC	.	.	Other	Intron	.
	APOC2	819	A	0.82	C	0.18	0.30	GGGACTTTGTACAGCMMAAGCACAGCAGCC	AAA	CAA	Nonsynonymous	Lys	Gln
	APOC3	1148	T	0.95	A	0.05	0.10	CTGGGGACTAAGAA WGTTTATGAACACCT	.	.	Other	Intron	.
	APOC3	1322	G	0.71	A	0.29	0.41	CACGGGCTTGAA TTRGGTCAGGTGGGGCC	.	.	Other	Intron	.
	APOC3	1468	A	0.97	C	0.03	0.06	ATACGGCTGAGCTCMGCCCTCTCTCAGAT	.	.	Other	Intron	.
	APOC3	1519	A	0.95	G	0.05	0.10	GGAGTGTGAACCCCTTTGTGAACTGGACA	.	.	Other	Intron	.
	APOC3	1637	T	0.96	A	0.04	0.07	GGCCCATGGA AAAA WTGTCCACCAACAAAA	.	.	Other	Intron	.
	APOC3	1722	A	0.84	G	0.16	0.27	AGGAAATGGGCCCGCCGCGCATGGCTCG	.	.	Other	Intron	.
	APOC3	1728	A	0.73	G	0.27	0.40	ATGGGGCCAGGGCGRGTGGCTCATGCTCG	.	.	Other	Intron	.

55

55	AVPR2EX1	114	A	0.97	T	0.03	0.06	TCA	TGGCGTCCACCWCTTCCGGTAAAGGCT	ACT	TCT	Nonsynonymous	Thr	Ser
	AVPR2EX2	109	G	0.98	A	0.02	0.04	ACCGGGACCCGCTCTCTAGCCGGGCGGA	CTG	CTA	CTA	Synonymous	Leu	Leu
	AVPR2EX2	129	C	0.85	T	0.15	0.25	CCGGGCGGAGCTGGYGCCTCTCTCCATAG	GCG	GTG	GTG	Nonsynonymous	Ala	Val
	AVPR2EX2	184	G	0.94	T	0.06	0.11	GGCTGTGTCTGGCKGCCCTAGCTGGGGG	GCG	GCT	GCT	Synonymous	Ala	Ala
	AVPR2EX2	444	C	0.87	T	0.13	0.23	CCGTCCATGCTGGYGTACCGCCATGGAA	GCG	GTG	GTG	Nonsynonymous	Ala	Val
	AVPR2EX3	112	C	0.95	T	0.05	0.10	TCTTTCAGCAGAGYGTGCTTCTAGAGCT	AGC	AGT	AGT	Synonymous	Ser	Ser
	AVPR2EX3	232	G	0.95	A	0.05	0.10	AAGGACACTTCATCTGTAGAGAGCTGTGG	TCG	TCA	TCA	Synonymous	Ser	Ser
	AVPR2EX3	252	T	0.97	C	0.03	0.06	AGCTGTGGGTGCTCTGTCTAGAGGCT	CTA	CTG	CTG	Other	3' UTR	
	BIR	46	A	0.50	G	0.50	0.50	GCGCCCTTGTGCTCTCTCATGTGTCTGGC	CTA	CTG	CTG	Synonymous	Leu	Leu
	BIR	1069	C	0.98	T	0.03	0.05	CAGGACTGGCTGGA YGCACAGCTCTAGGG				Other	3' UTR	
	BIR	1142	G	0.67	A	0.33	0.44	GGTGAGCCAGTCTCTRAATTGGGTTGGGAG				Other	3' UTR	
	BIR	1185	G	0.98	T	0.03	0.05	ATAACCCAGTACAGKTTCTCTGTGAGGCC				Other	3' UTR	
	BIR	1265	G	0.95	A	0.05	0.10	GGAGGCTGAGCTGARGCTGGCCCAAGCCTC				Other	3' UTR	
	BIR	1295	C	0.24	T	0.76	0.37	CACCAGGCCCTGGCYGGGCTACATACCAC				Other	3' UTR	
	BIR	1441	T	0.99	C	0.01	0.02	AGGGCCCGGGCYGAGGCGAGGGTCTAG	TCA	TCG	TCG	Synonymous	Ser	Ser
	BIR	1521	C	0.26	T	0.74	0.38	TGTGGGCACCTTGTGAYGGTGTGTCCAAACT	GTC	ATC	ATC	Nonsynonymous	Val	Ile
	BIR	1729	G	0.99	C	0.01	0.03	GGTGCCAGCTCGTASAGTGGGCTGTGGC	CTC	CTG	CTG	Synonymous	Leu	Leu
	BIR	1946	C	0.99	T	0.01	0.02	GCA TCAAGCAGAGGYGGCGGTGGCGCAGG	CGC	CAC	CAC	Nonsynonymous	Arg	His
	BIR	1960	G	0.75	A	0.25	0.38	CGGCCGTGGCGCAGRGCGATCACCCCATG	GCC	GCT	GCT	Synonymous	Ala	Ala
	BIR	2463	T	0.22	C	0.78	0.34	ACGGTACCTGGGCTYGGCAGGGTCTCTTG	AAG	GAG	GAG	Nonsynonymous	Lys	Glu
	BIR	2664	T	0.97	C	0.03	0.06	TGTCTGGCCTCACYCTCTGAGATAACTCC				Other	5' UTR	
	BIR	2894	T	0.99	G	0.01	0.03	TGGTGGTGGCACCKGTATATCCCACTTAC				Other	Promoter	
	BIR	2954	G	0.97	C	0.03	0.06	CCCGAGAGGGGAGSTTGCAGTGAAGCCAA				Other	Promoter	
	BIR	3174	C	0.70	T	0.30	0.42	TCCCGCTAAGAGCCYTTCTCCCGCCCCAG				Other	Promoter	
	BIR	369	G	0.97	A	0.03	0.06	CAACACTGCTCCAAARGGTCCAGGCGACGGG				Other	3' UTR	
	BIR	510	T	0.99	C	0.01	0.02	CCTTCTGGACAAAGYAGTGGCAGCCACT				Other	3' UTR	
	BIR	657	T	0.92	A	0.08	0.14	CACAGAGCCCTCACWGCACGAGGCCGATG				Other	3' UTR	
	BIR	981	G	0.98	A	0.03	0.05	TGGAGCCACAGACRCAAGAGCAGCAGCCC				Other	3' UTR	
	BKRB2EX1	55	T	0.77	G	0.23	0.35	GGTGGGACGGTGGKAGACGGTGGGGACAT				Other	5' UTR	
	BKRB2EX3	1513	T	0.93	C	0.07	0.13	ATCTCCAGGAGAACYGGCATCCAGCTTTG				Other	3' UTR	
	BKRB2EX3	1833	G	0.95	A	0.05	0.10	ACTCAAGTGGGAACRACACTGGGCACCTGCCA				Other	3' UTR	

5	10	15	20	25	30	35	40	45	50	55
Thr	Thr	ACA	ACG	AAGGAGATCCAGACRAGAGAGGGCCAC	0.13	0.07	A	G	747	BKR2EX3
Leu	Arg	CTT	CGT	TTTCTGGGAGGTCTTCCACCCCGTGG	0.02	0.01	T	G	343	BNPEX1
Intron	Intron			TGAGGCTTGGACGSCCCCATTCATTGCAG	0.07	0.03	G	C	15	BNPEX2
Met	Met	TTG	ATG	GTGGCACCGCAAAWTGGTCTCTACACC	0.02	0.01	T	A	174	UNPEX2
Leu	Nonsynonymous		CGC	ATTGCAGGAGCAGCAACCATTTGCAGG	0.06	0.03	A	G	37	DNPEX2
Ile	Nonsynonymous	CAC		AGAACTGAAGCAAAARGAGTATCTGGATGT	0.10	0.05	A	G	424	BRSEX1
Promoter	Other			GTGCCATCTATATTMCTTATGCTGTGATC	0.06	0.03	C	A	730	BRSEX1
Pro	Thr	CCT	ACT	CTAACTTGTGTGCGWGTGGATGCAACTCA	0.10	0.05	T	A	879	BRSEX1
Pro	Pro	CCT	CCA	GCTCTACCTGAGGCWATATTTTCAAAATGT	0.11	0.06	A	T	144	BRSEX2
Ala	Synonymous	GCA	GCT	CTCCAATGCCATCCWGAAGACTTGTGTAA	0.04	0.02	A	T	80	DRSEX2
Gln	Nonsynonymous	CAG	CTG	GCCATGCAATTCAATYTTCCACCATTTCTC	0.12	0.06	C	T	173	BRSEX3
Ile	Synonymous	ATC	ATT	CCCCAGTCACAGGCKCTGGAGGCAAGAG	0.05	0.02	G	T	1063	CAL/CGRPEX1+2
	Other			GTCCGATCAGGCACRCCGCTTGGAGCCCA	0.24	0.14	A	G	940	CAL/CGRPEX1+2
5' UTR	Other			CTGCACTGGTGCACTATGTCAGATG	0.15	0.08	A	G	112	CAL/CGRPEX3
5' UTR	Nonsynonymous	AAC	GAC	GTGCAGACTATGTCAGATGAAAGGCCAG	0.02	0.01	T	G	120	CAL/CGRPEX3
Asn	Asp	GTT	GTG	TGTTTCCCTGCGAGMCTGGACAGCCCGAG	0.16	0.09	A	C	30	CAL/CGRPEX4
Val	Val	AGA	AGC	ATGTGGTTTAAAWATCCATAAGGGAAG	0.48	0.41	T	A	309	CAL/CGRPEX5
Arg	Ser			CAGACCAAGAAATAYAGATCTTGTTTATT	0.34	0.22	T	C	433	CAL/CGRPEX5
	3' UTR			AAAGAGCAAGTGAGRTAATAGATGTTAAG	0.16	0.09	A	G	719	CAL/CGRPEX5
	3' UTR			TTCCTTCTGGGAGWTATAAACCCCAAGA	0.24	0.14	A	T	158	CHYEX1
	3' UTR			TCTAGGGGAACCTCYGATCAGAAACAGCC	0.42	0.30	C	T	65	CHYEX1
	3' UTR			TTGTAACTTCCAACSGTCCCTCAAAATTT	0.09	0.05	C	G	107	CHYEX2
	3' UTR			GCTGACGGCTGCTCTTTGTCAGGAAGGT	0.14	0.08	G	A	168	CHYEX2
	3' UTR			CTTCTTCTCACAACAGGCTTATAACAG	0.17	0.09	G	A	26	CHYEX3
	3' UTR			CTCCCTTCCCATCMCAATTCAACTTGT	0.15	0.08	C	A	83	CHYEX4
Arg	Gly	CGT	GGT	TCCCTCAGCCACAAYCCTAAGCTCCAGA	0.19	0.11	T	C	274	CHYEX5
Arg	Ile	CGT	CAT	CTCTGCCACCTTGSTTCTCGCCCTCCATC	0.49	0.44	C	G	33	CLCNKBEX10
Intron	Intron	TCC	TCA	GGAGGAGCTGCTATYGGGCGCTCTTTGG	0.12	0.06	T	C	12	CLCNKBEX13
Ser	Ser			ACTGCCAAAGACAYGCCACTGGAGGAGG	0.11	0.06	T	C	64	CLCNKBEX15
Leu	Val	CTT	GTT	GCCAAGGACAGCCRCCTGGAGGAGTGGT	0.45	0.66	G	A	68	CLCNKBEX15
Ile	Ile	ATT	ATC	CCCTTTGTGACGTGCGGGGCAGAGCTG	0.08	0.04	T	C	51	CLCNKBEX18
Met	Thr	ATG	ACG							
Pro	Pro	CCG	CCA							
Trp	Ser	TGG	TGG							

	5	10	15	20	25	30	35	40	45	50	55
CLCNKBEX3	Arg										
CLCNKBEX3	Ser	Nonsynonymous	CGC	AGC	GGGAGATTGGGACACGCCACCTGCTCCGG	0.11	0.06	C	0.94	A	34
CLCNKBEX4	Ser	Synonymous	TCA	TCG	GTCTCTTCTCTTCRGGCTTCTCTCAGAG	0.13	0.07	A	0.93	G	96
CLCNKBEX4	Val	Nonsynonymous	CTG	GTG	CTGGAATCCCGAGSTGAAGACCATGTTG	0.46	0.65	C	0.35	G	19
CLCNKBEX4	Asn	Nonsynonymous	CAC	AAC	ACCTGGATATCAAGMACCTTTGGGGCCAAA	0.15	0.08	C	0.92	A	70
CLCNKBEX7	Ala	Synonymous	GCA	CCG	TTCCGGCTCTCGGCKGTCTTCAACAGCGA	0.30	0.11	A	0.89	G	108
CNPEX1	5' UTR	Other	.	.	GCAGGCCAACTTTMTGCCCTGTATGACTT	0.16	0.09	A	0.91	C	1018
CNPEX1	5' UTR	Other	.	.	GCCTTCACGCCCTGGKGACAGCCACTGCAC	0.04	0.02	T	0.98	G	144
CNPEX1	5' UTR	Other	.	.	GCAGCACTGGGACCSCTGCTCGCCCTGCAG	0.05	0.02	G	0.98	C	1457
CNPEX1	5' UTR	Other	.	.	ATTGTTCCACAGARGGAGTTTCAACAGCG	0.16	0.09	A	0.91	G	578
CNPEX1	5' UTR	Other	.	.	GGGAGTTTCAACAGCRGAGTCAGACCCCGG	0.11	0.06	A	0.94	G	592
CNPEX2	3' UTR	Other	.	.	AACATCCAGCTCAGACATTGACAGTCA	0.14	0.08	A	0.92	T	1171
CNPEX2	Gln		CAG	CGG	GGACACCAAGTCGCRGGCAGCGTGGGCTC	0.08	0.04	A	0.96	G	139
CNPEX2		Nonsynonymous	.	.	CCCCCGCCCAAGCRGCCCTTCGAGGGCGC	0.10	0.05	G	0.95	A	357
CNPEX2		Other	.	GGT	GCTCGGGCGCGGCGYCAAGAAGGGCGA	0.03	0.02	C	0.98	T	41
COX1	Gly	Synonymous	GGC	AAA	TTTCTGCAAGTCTGARAATTGACCCAGAGC	0.02	0.01	G	0.99	A	1063
COX1	Lys	Nonsynonymous	AGA	ATC	ACATGGACCAACACRCTCTGTCATGTGGCT	0.02	0.01	G	0.99	A	1314
COX1	Val	Nonsynonymous	GTC	AAG	TCAAATGAGTACCCGCRAGAGGTTTGGCATG	0.02	0.01	G	0.99	A	1386
COX1	Glu	Nonsynonymous	GAG	GTA	CCTTCCAGGAOCTCTTAGGAGAGAAGGAG	0.03	0.02	A	0.98	G	1428
COX1	Ile	Nonsynonymous	ATA	.	GGTGAAGTTTGGGGYTGACATTTAGAACT	0.08	0.04	C	0.96	T	1906
COX1		Other	.	.	ATTAATCTGGAATAATGCTGCTTTTAT	0.02	0.01	C	0.99	T	1948
COX1		Other	.	.	GTCCTCCAGAAATACGGGTTCTTAGTTGA	0.02	0.01	G	0.99	T	2037
COX1		Other	.	.	TGCCACCTTCATCCRAGAGATGCTCATGC	0.02	0.01	A	0.99	G	310
COX1	Gln	Nonsynonymous	CAA	CGA	TTCAAAACTTCTGGMAAGATGGGTCTGG	0.16	0.09	A	0.91	C	626
COX1	Gly	Synonymous	GGA	GGC	TTTATGGAGACAAATMTGGAGCGTCAGTAT	0.03	0.02	A	0.98	C	696
COX1	Met	Nonsynonymous	ATG	CTG	GACCTGCTGAAGGCGYAGCACCCCACTG	0.04	0.02	C	0.98	T	938
COX1	Ala	Synonymous	GCC	GCT	CGATTTTCTCATTTSCGTGGGTAAAAAC	0.27	0.16	G	0.84	C	186
COX2EX1		Other	.	.	CGGACCAATTGTCAKACUACTTCGAGTGA	0.27	0.16	G	0.84	T	358
COX2EX1		Other	.	.	AACCATGGTAGAAGYTGAGACCACTTCT	0.12	0.06	C	0.94	T	156
COX2EX10	Ala	Nonsynonymous	GCT	GTT	GCAAGTCTCTCCCGMTCCGGACTAGTGA	0.05	0.03	A	0.98	C	379
COX2EX10	Arg	Synonymous	CGA	CCG	AAAGTACTTTTGGTYATTTTTCGTGTATC	0.50	0.49	C	0.51	T	866
COX2EX10	3' UTR	Other	.	.	CATCGATGCTGTGGRCGTGTATCTGCCCC	0.02	0.01	G	0.99	A	87
COX2EX10	Glu	Nonsynonymous	GGG	GAG							

55	COX2EX10	937	G	0.83	A	0.17	0.28	ATTAACATTACCACTAAATTCATGCTA	GTG	GTC	Other	3' UTR	Val
	COX2EX3	166	G	0.93	C	0.07	0.13	ATTATGAGTTATGTSTTGACATGTAAGTA	GAT	GAC	Synonymous	Val	Val
	COX2EX7	206	T	0.96	C	0.04	0.07	AACAGAGTATGCCAYGTGCTTAAACAGGA	CAT	CAC	Synonymous	Asp	Asp
	COX2EX8	268	T	0.95	C	0.05	0.10	ATATTGCTGGAAACAYGGAATTAACCAATT			Synonymous	His	His
	CYP11B1EX1	351	T	0.97	C	0.03	0.06	TGACGTGATCCCTCYGGAAGGCAAGGCAC	CCC	CCG	Other	Promoter	
	CYP11B1EX1	525	C	0.99	G	0.01	0.03	AGGACAGTCTGCCSTTTGAAAGCCATGCC	CGG	CAG	Synonymous	Pro	Pro
	CYP11B1EX1	542	G	0.97	A	0.03	0.06	TGAAGCCATGCCRCGCTCCAGGCAACA	GAC	CAC	Nonsynonymous	Arg	Gln
	CYP11B1EX1	601	G	0.97	C	0.03	0.06	AGCAGGTTATGAGSACCTGCACCTGGAA	CTG	TTG	Nonsynonymous	Asp	His
	CYP11B1EX2	184	C	0.99	T	0.01	0.03	GTGCGTGTCTCTTGTTGTAAGCGCGGAGC	GAT	GAC	Synonymous	Leu	Leu
	CYP11B1EX2	188	A	0.96	G	0.04	0.07	CGTCTTCTGCTGTGTRAGCGCGGAGCTGAG	ATC	ATC	Other	Intron	
	CYP11B1EX2	36	T	0.46	C	0.54	0.50	CCCCACAGGTACGAYTTGGAGGAGCAGG	GAT	GAC	Synonymous	Asp	Asp
	CYP11B1EX2	78	C	0.96	T	0.04	0.08	ATGCTGCCGGAGGAYGTGGAGAAGCTGCA	ATG	ATC	Synonymous	Met	Ile
	CYP11B1EX3	114	G	0.99	C	0.01	0.02	AGTTCTCCCGATSGTGGATGCAATGGC	TTT	TTG	Nonsynonymous	Thr	Ile
	CYP11B1EX4	177	C	0.98	T	0.02	0.04	CTGTCTCGCTGGAYCAGCCCAAGGTGT	GCG	GCA	Nonsynonymous	Phe	Leu
	CYP11B1EX4	205	T	0.92	G	0.08	0.15	TGGAAGGAGCACCTTKGAGCCCTGGGACTG	CTC	GTC	Other	Intron	
	CYP11B1EX4	247	C	0.91	G	0.09	0.16	GGTGAGGCCAGGAGSCCGGCAAGTCTAT	TAT	TAC	Synonymous	Ala	Ala
	CYP11B1EX5	103	G	0.97	A	0.03	0.06	ACCAGCATCTGGCGAGCTCTCTGTTGAA	CTG	GTC	Nonsynonymous	Leu	Val
	CYP11B1EX5	107	C	0.84	G	0.16	0.26	GCATCGTGGCGGAGSTCCTGTTGAAATCG			Other	Intron	
	CYP11B1EX5	16	C	0.58	T	0.42	0.49	TGAGGGCTGCCCTCCYGCCTCCCGGATAGG	AUC	AAC	Synonymous	Tyr	Tyr
	CYP11B1EX5	55	T	0.97	C	0.03	0.06	ATCCAGAAAATCTAYCAGGAACCTGCCCTT	CTG	TTG	Nonsynonymous	Ser	Asn
	CYP11B1EX5	72	G	0.99	A	0.01	0.03	GGAACTGGCTTCARCCGCCCTCAACAGT	TAC	CAC	Synonymous	Leu	Leu
	CYP11B1EX7	52	C	0.99	T	0.01	0.03	CTGTGUTCTGTTTCTGAGCGAGTGCGG			Nonsynonymous	Tyr	His
	CYP11B1EX8	144	T	0.96	C	0.04	0.08	CCGCAGGAACCTTCYACCAGCTGCCCTTT			Other	Intron	
	CYP11B1EX9	16	G	0.96	C	0.04	0.07	CCAGATGGAAACCCCGCTTCTGTCTTAGG			Other	3' UTR	
	CYP11B1EX9	274	T	0.91	C	0.09	0.16	AGCCCCAGCACAAAYGGAATCTCCGAGGG			Other	3' UTR	
	CYP11B1EX9	350	T	0.88	G	0.12	0.21	GCTGGGGAAGATCTKGCTGACCTTGTCC			Other	3' UTR	
	CYP11B1EX9	459	G	0.72	A	0.28	0.40	CTCTGTGTGGCCATRCAGGGTGCTGTGG			Other	3' UTR	
	CYP11B1EX9	592	A	0.93	C	0.07	0.13	TCTAGAGTCCAGTCMAGTTCCCTCTGCA	ACC	ACT	Other	3' UTR	
	CYP11B1EX9	62	C	0.99	T	0.01	0.03	GTGGAGACACTAACCYCAAGAGGACATAA			Synonymous	Thr	Thr
	CYP11B1EX9	657	G	0.66	A	0.34	0.45	CTCTGAAAGTTGTCTCCCTCGGAATAGGGT			Other	3' UTR	
	CYP11B1EX9	786	A	0.87	G	0.13	0.22	ATCGTGTGAGGCTCTGTGCCCTGGGCTCA			Other	3' UTR	

55	CYP11B1EX9	835	C	0.77	T	0.23	0.35	GTTCAGAGAGTGGYGTGGGTCTCTCTC	10	Other	3' UTR	Phe
	CYP11B1EX9	879	A	0.68	G	0.32	0.44	CTGGGAAAGTCCCRAGGATGCTGTCAAG		Other	3' UTR	Arg
	CYP11B2EX1	163	A	0.97	G	0.03	0.06	TCCTGGGTGAGATARAAGGATTGGGCTG		Other	Promoter	
	CYP11B2EX3	138	C	0.66	T	0.34	0.45	GTGGCAGGAGCTTTCACAGCCCTGAA	15	Synonymous	TTT	Phe
	CYP11B2EX3	152	A	0.71	G	0.29	0.41	CTCCAGGCCCTCARGAAGAAGGTGCTGC		Nonsynonymous	AGG	Lys
	CYP11B2EX3	20	G	0.98	C	0.03	0.05	CAAGCTCTGCCCTGSCCTCTGTAGGAATG		Other	Intron	
	CYP11B2EX3	243	G	0.97	A	0.03	0.06	GGTGTGGGCCATGCRGAAGGTCCAGCCC		Other	Intron	
	CYP11B2EX4	177	T	0.96	C	0.04	0.07	CCTGTCTCGTGGAYCAGCCCCAAGGTGT		Nonsynonymous	ACC	Thr
	CYP11B2EX4	250	G	0.98	A	0.02	0.04	QAGGCCAGGACCCGCGCAGTGTCTATGGG		Other	Intron	
	CYP11B2EX4	99	A	0.94	C	0.06	0.12	TTCTGCCAGCCTGAMCTTCCTCCATGCC		Nonsynonymous	ACC	Thr
	CYP11B2EX5	103	G	0.05	A	0.95	0.10	ACAGGATCGTGGCRGAGCTCTCTGTGAA		Synonymous	GCA	Ala
	CYP11B2EX5	121	G	0.78	A	0.23	0.35	CTCCTGTTGAAGGCRGAACCTGTACTAGA		Synonymous	GCA	Ala
	CYP11B2EX5	55	C	0.99	T	0.01	0.02	ATCCAGAAATCTAYCAGGAACCTGGCCTT		Synonymous	TAT	Tyr
	CYP11B2EX5	72	A	0.99	G	0.01	0.02	GGACTGGCCTTCARCCGCCCTCAACACT		Nonsynonymous	AAC	Ser
	CYP11B2EX6	195	A	0.84	C	0.16	0.27	TCAAGGAGACCTTGMGGTGGGTGCTGGCT		Synonymous	CGG	Arg
	CYP11B2EX6	91	T	0.38	C	0.63	0.47	CGACUTGCACAGAYCCTGCCCCAGGAGA		Nonsynonymous	ACC	Thr
	CYP11B2EX7	52	T	0.99	C	0.01	0.03	CTGTGGTCTGTTTGTGGACGAGTGGTG		Synonymous	CTG	Leu
	CYP11B2EX7	56	A	0.97	T	0.03	0.06	GGTCTGTTTGTGGWCGAGTGGTGAGCT		Synonymous	GTG	Val
	CYP11B2EX7	65	T	0.91	C	0.09	0.17	TTGGAGCGAGTGGYGAGCTCAGACTTGG		Nonsynonymous	GCG	Ala
	CYP11B2EX7	78	G	0.82	A	0.18	0.30	GTGAGCTCAGACTTGTGCTTCAGAACTA		Synonymous	TTA	Leu
	CYP11B2EX8	132	G	0.97	A	0.03	0.06	ACATCAGGGGCTCCGCGAGGAACCTCCAC		Nonsynonymous	AAC	Ser
	CYP11B2EX8	18	C	0.99	T	0.01	0.02	TGATCCCTGCTCTGYACCCGTCGCGAGACA		Other	Intron	
	CYP11B2EX8	182	C	0.98	T	0.02	0.04	ATGCGCCAGTGCCTYGGGCGGCGCCTGGC		Synonymous	CTT	Leu
	CYP11B2EX8	37	T	0.99	A	0.01	0.02	TCCGACAGACATTGGWACAGGTTTTCCTCT		Nonsynonymous	GAA	Glu
	CYP11B2EX9	224	G	0.89	A	0.11	0.20	GTCTTCTCTCCACRTGCACAGCTTCCTG		Other	3' UTR	
	CYP11B2EX9	90	T	0.99	G	0.01	0.03	AGATGGTCTACAGCKTCAATTGAGGCCT		Nonsynonymous	GTC	Val
	DBHEX1	152	G	0.99	A	0.01	0.02	GGGCCAGCCTGCCRCGCCCCAGCATCCGG		Other	Phe	
	DBHEX3	153	G	0.92	A	0.08	0.14	AGTTGCCCTCAGACRCGTGCACCATGGAG		Nonsynonymous	ACC	Thr
	DBHEX3	239	G	0.97	C	0.03	0.06	AAGGAGCTTCCAASGGCTCTCTCGGCA		Nonsynonymous	AAC	Asn
	DBHEX3	257	C	0.98	T	0.03	0.05	TTCTCTGGCACCA YATTATCAAGGTACG		Synonymous	CAT	His
	DBHEX3	63	G	0.96	C	0.04	0.08	CGTCCGGTCACTGSAAGGCCATCAACGGC		Nonsynonymous	CAG	Gln

SDOCID: <EP_____.0955382A2_I_>

55	EDNRBEX4	39	G	0.82	A	0.18	0.29	AAAGATTGGTGGCTTTCAGTTTCTATTT	CTG	CTA	Synonymous	Leu	Leu
	ELAM1EX1	143	A	0.93	G	0.07	0.13	TCTTTGACCTAAATRATGAAGTCTTAAA	.	.	Other	Promoter	.
	ELAM1EX1	209	T	0.97	G	0.03	0.06	TTAATTGCACTAGTGKCTTTGGCCAAAT	.	.	Other	Promoter	.
	ELAM1EX10	107	C	0.91	T	0.09	0.16	CATTAGCAACATTTTTCCTCTGGCTTCGG	CTC	TTC	Nonsynonymous	Leu	Phe
	ELAM1EX12	54	T	0.96	C	0.04	0.08	AGCCTTGAATCAGAYGGAAGCTACCAAAA	GAT	GAC	Synonymous	Asp	Asp
	ELAM1EX13	1004	T	0.98	G	0.02	0.04	CAGAAATATGTGCTKTCACCATGAAAAA	.	.	Other	3' UTR	.
	ELAM1EX13	1158	G	0.18	A	0.82	0.29	GATGTTTGTCAGATRTGATATGTAAACAT	.	.	Other	3' UTR	.
	ELAM1EX13	1549	G	0.39	A	0.61	0.47	TGAACACTGGCAACRACAAAGCCACAGT	.	.	Other	3' UTR	.
	ELAM1EX13	967	T	0.97	C	0.03	0.06	ACTGAATGGAAGGTYTGTATATTTGTCAGA	.	.	Other	3' UTR	.
	ELAM1EX2	382	C	0.98	G	0.02	0.04	AATGATGAGAGGTGAGCAAGAAAGAGCT	.	.	Nonsynonymous	Cys	Trp
	ELAM1EX3	152	T	0.95	C	0.05	0.10	GTAAGTCTGGTTCTTGGCTCTTTCTTCAC	TGC	TGG	Nonsynonymous	Intron	.
	ELAM1EX3	53	A	0.97	C	0.03	0.06	CCAATACATCCTGCMGTGGCCACCGTGAA	AGT	CGT	Other	Ser	Arg
	ELAM1EX5	197	G	0.95	A	0.05	0.10	GGAAITGGGACAAACRAGAAAGCCAACTGT	GAG	AAG	Nonsynonymous	Glu	Lys
	ELAM1EX5	55	T	0.97	C	0.03	0.06	GATGCTGTGACAAAYCCAGGCCAAATGGTT	AAT	AAC	Synonymous	Asn	Asn
	ELAM1EX7	199	C	0.94	T	0.06	0.12	GGGAGTGGGACAAAYGAGAAAGCCACATG	AAC	AAT	Synonymous	Asn	Asn
	ELAM1EX7	200	G	0.94	C	0.06	0.11	GGAGTGGGACAAACRAGAAAGCCCACTGT	GAG	CAG	Nonsynonymous	Glu	Gln
	ELAM1EX8	152	C	0.96	T	0.04	0.08	AGGATTTGAAATTATATGATCAACTCAA	CAT	TAT	Nonsynonymous	His	Tyr
	ELAM1EX8	22	T	0.91	C	0.09	0.16	AGTGCTCTCTGTYGTTCCAGCTGTGAG	.	.	Other	Intron	.
	ENDOTHELIN2	440	T	0.97	C	0.03	0.07	CCCTTGCAGACGTGYTCCAGACTGGCAAG	TTC	CTC	Nonsynonymous	Phe	Leu
	ENDOTHELIN2	556	G	0.99	A	0.01	0.02	ATGCGGAGCCTCGRTCCACACATTCACAG	CGG	CGA	Synonymous	Arg	Arg
	ENDOTHELIN2	976	A	0.84	G	0.16	0.27	AGCCAGCCCTGGAGKCTGGATGGCTCCCC	.	.	Other	3' UTR	.
	ET1EX3	114	G	0.88	A	0.12	0.21	GCAACAGACCGTGARAATAGATGCCAATG	GAG	GAA	Synonymous	Glu	Glu
	ET1EX5	90	G	0.69	T	0.31	0.43	AAGCTGAAAGGCAAKCCCTCCAGAGAGCG	AAG	AAT	Nonsynonymous	Lys	Asn
	GALNREX1	1052	G	0.94	T	0.06	0.11	CTGCCACCTGGGTCTGGCGGCTTTCAT	GTG	GTT	Synonymous	Val	Val
	GALNREX1	325	C	0.98	G	0.02	0.04	GGTGACGACCGCAGCGCTCCGGGAGCCA	.	.	Other	Promoter	.
	GALNREX1	327	G	0.81	C	0.19	0.30	TGCAGCAGCGCCCTCCGGGAGCCAGG	.	.	Other	Promoter	.
	GALNREX1	553	G	0.49	C	0.51	0.50	TCTCTCAGAAAGTCCGGGCCAAGACCGG	.	.	Other	Promoter	.
	GALNREX1	887	C	0.49	T	0.51	0.50	ATCTTCGGCGCTGGYGTGCTGGGCAACAG	GGC	GGT	Synonymous	Gly	Gly
	GALNREX3	298	A	0.68	G	0.32	0.43	TGATAAAGAAARTAAAGTCGAATAG	AAT	AGT	Nonsynonymous	Asn	Ser
	GALNREX3	322	C	0.98	T	0.02	0.04	AATAGACACCCACATCAACCAATTGTA	CCA	CTA	Nonsynonymous	Pro	Leu
	GALNREX3	388	T	0.98	C	0.02	0.04	AGTTCCCATATAAGYGGACCCAGACACAGA	.	.	Other	3' UTR	.

55	50	45	40	35	30	25	20	15	10	5	
GALNREX3	418	C	0.97	G	0.03	0.06	ACAAACAGAAATGAAGTACTAGTAAACGATGCT	CTG	CTA	Leu	3' UTR
GALNREX3	523	G	0.98	T	0.02	0.04	TAGGAAATTCCTAGKTCCTAGTGAGAAATTA	TTC	TTT	Phe	3' UTR
GALNREX3	650	C	0.94	T	0.06	0.11	TCCATATATATGTTAACTCTTCATAGAT	CCT	GCT	Pro	3' UTR
GALNREX3	799	A	0.84	G	0.16	0.26	ATGTATTTTAAATRGATCATGGACACA	GCC	GCG	Ala	3' UTR
GGREX1	125	G	0.98	A	0.02	0.04	CTGCTGTGCTCTTCTCTGCTGGCTGCCA	TTC	TTT	Leu	
GGREX11	57	C	0.91	T	0.09	0.16	CACGAAAGTGTCTTTCGCTTCGTGACGGA	CCT	GCT	Phe	
GGREX4	68	C	0.98	G	0.02	0.05	GACCCGGGGGAGSCCTTTCGCTGATGCTC	GCC	GCG	Ala	
GGREX5	71	C	0.83	G	0.17	0.28	CTGTCCCTGGGGGCTCTCTCTCTCTCTCT	TTC	TGC	Ala	
GGREX9	29	T	0.93	G	0.07	0.12	TGACAAATGGGCTCTCTCTCTCTCTCTCT	GTC	ATC	Cys	
GHIEX4	144	G	0.98	A	0.02	0.03	CCTCTGACAGCAACRTCTATGACCTCCTA	AGG	TGG	Ile	
GHIEX3	126	A	0.94	T	0.06	0.12	CAACACCTTCCAAACWGGGTGAAACCCAG	AAC	AAT	Trp	
GIPREX2	72	C	0.93	G	0.08	0.14	CTTCCGCCCTCTCAGATGACTACTCTCTC	CTC	GTC	Asn	
GIPREX7	51	C	0.79	T	0.21	0.33	CAITGGCACTAGAAATATATCCACATCAA			Val	
GIPREX8	180	C	0.98	G	0.03	0.05	GCTACTACCTGCTCTCTCTCTCTCTCTCT			Promoter	
GLUT2EX1	137	C	0.32	A	0.68	0.44	CCACAGCACTAATTTCTCTCTCTCTCTCT			Promoter	
GLUT2EX1	164	T	0.31	C	0.69	0.43	AGTCCAGTGTGCTCTCTCTCTCTCTCTCT			Promoter	
GLUT2EX1	237	T	0.96	C	0.04	0.07	AAAGATTCTCTCTCTCTCTCTCTCTCTCT			Promoter	
GLUT2EX1	242	G	0.34	A	0.66	0.45	TTTCTCTCTCTCTCTCTCTCTCTCTCTCT			Promoter	
GLUT2EX10	161	G	0.99	A	0.01	0.02	GAATTCACAAAGAAAGAGTGGCTCAGCCCA	AAG	AAA	Lys	
GLUT2EX10	87	C	0.99	G	0.01	0.03	TCCTGGCTTTTACCTCTCTCTCTCTCTCT	CTG	GTC	Val	
GLUT2EX10	92	T	0.35	C	0.65	0.46	GCCTTTACCTCTCTCTCTCTCTCTCTCT	TTT	TTC	Phe	
GLUT2EX3	250	C	0.87	T	0.13	0.23	AGTTGGTGAATGATCTCTCTCTCTCTCT	ACT	ATT	Ile	
GLUT2EX4A	153	T	0.96	C	0.04	0.08	TCAGGACTATATGCGGTAAAGTCTCACAC	TGT	TGC	Cys	
GLUT2EX4A	162	T	0.83	A	0.17	0.28	TATTTGGTAAAGTCTCTCTCTCTCTCTCT			Intron	
GLUT2EX4A	164	A	0.94	T	0.06	0.11	TTGTGGTAAAGTCTCTCTCTCTCTCTCT			Intron	
GLUT2EX4B	127	A	0.28	G	0.72	0.40	CTGCCCATCTCTCTCTCTCTCTCTCTCT	ACA	ACG	Thr	
GLUT2EX5	78	C	0.93	T	0.08	0.14	ATCTGTGGCAGATCTCTCTCTCTCTCTCT	CTG	TTG	Leu	
GLUT2EX6	15	T	0.10	C	0.90	0.18	TGTTTCAACCTGATATTTCTCTCTCTCT			Intron	
GLUT2EX8	21	T	0.88	C	0.12	0.21	TAATTTCTTTAAAAATGCTCTCTCTCTCT			Intron	
GLUT2EX8	38	T	0.95	C	0.05	0.10	TCCTAGGATATCTCTCTCTCTCTCTCTCT	CTT	CTC	Leu	
GLUT4EX1	1002	A	0.99	C	0.01	0.02	CGTTGTGGGAAACGGMAATTTCTCTCTCT			Promoter	

5

10

15

20

25

30

35

40

45

50

55

GLUT4EX1	1051	C	0.74	T	0.26	0.38	ACCATGTCGGGACACCTTTAAGGCGTCAT	Other	Promoter
GLUT4EX1	1228	A	0.86	T	0.14	0.24	TTCTCAGGCCGCTGGWGTTCCTCCCGGGGCA	Other	Promoter
GLUT4EX1	1632	A	0.71	C	0.29	0.42	CAGCCCCGCTCCACMAGATCCCGGGGAGC	Other	5' UTR
GLUT4EX1	1662	A	0.64	G	0.36	0.46	CCACTGCTCTCCGGRTCTCTGGCTTGTGG	Other	5' UTR
GLUT4EX1	1683	G	0.98	C	0.02	0.05	GCTTGTGGCTGTGGSTGCCATCGGGCCCG	Other	5' UTR
GLUT4EX1	1691	G	0.94	A	0.06	0.11	CTGTGGGTCCCATCRGGCCGCCCTCCGA	Other	5' UTR
GLUT4EX1	368	G	0.96	A	0.04	0.07	ACAGAGGAATCGARCCCTGACTTTTACCA	Other	Promoter
GLUT4EX1	560	A	0.90	G	0.10	0.19	CCGGAAGGGAGAGATAGTGGGTTGAGA	Other	Promoter
GLUT4EX1	615	G	0.93	A	0.07	0.14	TCGCTCGCCCTCCARGTGGCAGCACAAAC	Other	Promoter
GLUT4EX1	91	C	0.95	T	0.05	0.10	CAGGAGGTTTGTCTTCTGAAAGGGA	Other	Promoter
GLUT4EX1	966	C	0.95	A	0.05	0.10	CTGAAGACAGGAGCMAAGCAGCCCGGCCA	Other	Promoter
GLUT4EX10	19	C	0.94	G	0.06	0.11	TCCACCTCCCTGTSTGGCCCTAGGAGC	Other	Intron
GLUT4EX11	1005	G	0.83	A	0.17	0.28	GTGCTGGGATTACARGCGTGAGCCACCCGC	Other	3' UTR
GLUT4EX11	1099	A	0.95	C	0.05	0.10	GAAAGTATGTCCTGTCGTCGCAAGATG	Other	3' UTR
GLUT4EX11	791	T	0.93	C	0.07	0.14	CGAGTCAGTGGCGYGATCTTCTTCACT	Other	3' UTR
GLUT4EX11	827	C	0.90	T	0.10	0.18	GTCTCCAGGTTCAYGCCATTTCTCTGCC	Other	3' UTR
GLUT4EX11	872	G	0.79	A	0.21	0.33	CTGGGACTACAGGRCATGCCACCAACACC	Other	3' UTR
GLUT4EX11	874	A	0.92	C	0.08	0.15	GGGACTACAGGCGMTGCCACCAACCTTG	Other	3' UTR
GLUT4EX11	884	A	0.73	G	0.27	0.40	CGCATGCCACCAACRCCTGGCTAAITTTAT	Other	3' UTR
GLUT4EX11	897	A	0.89	T	0.11	0.20	CACCTGGCTAAITTTTGTATTTTGTAG	Other	3' UTR
GLUT4EX11	930	A	0.88	G	0.13	0.22	TACGCGTTTCACRCRTGTGTAGCCAGAAATG	Other	3' UTR
GLUT4EX11	935	A	0.86	G	0.14	0.24	GTTTTCACCATGTTTGGCCAGAAATGGTCTC	Other	3' UTR
GLUT4EX11	941	A	0.49	G	0.51	0.50	ACCATGTTAGCCAGRATGGTCTCGATCTC	Other	3' UTR
GLUT4EX11	963	C	0.96	T	0.04	0.07	CGATCTCTGACCTTGATCTGCTGCCC	Other	3' UTR
GLUT4EX3	112	C	0.90	G	0.10	0.17	TCCAGGACCCCTCASCACCCCTCTGGGCCC	Nonsynonymous	Thr
GLUT4EX4	96	C	0.71	T	0.29	0.42	ATGGGCTGGCCAAVCGTCTGCTGCTCTCTA	Synonymous	Asn
GLUT4EX7	19	C	0.95	T	0.05	0.09	TCAGGCTGACCTTTCCTTCTTCCAGGTCT	Other	Intron
GLUT4EX7	227	G	0.99	C	0.01	0.02	ATGCTGTATGTGTGAGCAGCAGCCTCCAGGC	Other	Intron
GLUT5EX1	184	G	0.95	A	0.05	0.10	AAAAGGAGGTGAGRCACATCTGCCCCCTTC	Other	5' UTR
GNB3EX1	184	A	0.89	C	0.11	0.20	GACAGATGGGGAACMCTGTGCTTCCCTGA	Other	Promoter
GNB3EX1	201	A	0.63	G	0.37	0.47	GTGCTCCCTGAAACRGAATAATGCAAGGGA	Other	Promoter

JSDOCID: <EP_____0955382A2_I_>

55	IIAPTEX1	375	A	0.89	G	0.11	0.19	ATTACGAATGGACRAGGGGAGAGTGA	GTG	Other	5' UTR	Val
	HAPEX4	34	A	0.53	G	0.47	0.50	TTTGTTTCAGGAGTRTACACCTTAAATGA	GTG	Synonymous	Val	Val
	HAPEX6	34	G	0.66	A	0.34	0.45	TTTGTTTCAGGAGTRTACACCTTAAACAA	GTG	Synonymous	Val	Val
	HAPEX7	1331	C	0.98	T	0.02	0.03	CATGCTGTTGCCCTCYTCAAAGTGAATTAG		Other	3' UTR	
	HAPEX7	367	G	0.98	A	0.02	0.03	GTGCTGTTAATGARAGAGTGATGCCCAT	GAG	Synonymous	Glu	Glu
	HAPEX7	610	T	0.98	C	0.02	0.04	CACACCTICTGTGTCYGGCA TGCTAAAGTA	GCT	Synonymous	Ala	Ala
	HAPEX7	673	C	0.98	T	0.02	0.03	GCCTTGCCGTTTCA YGACCTGGAGGAGGA	CAC	Synonymous	His	His
	HSDIKEX2	232	C	0.95	A	0.05	0.10	ACCAAGGCCACACMACCACCACCGGTCA	ACC	Synonymous	Thr	Thr
	HSDIKEX3	139	G	0.83	A	0.18	0.29	AATTCTTTGGCGRCTCGAGCTGACCCAA	GCG	Synonymous	Ala	Ala
	HSDIKEX5	951	T	0.96	A	0.04	0.08	ACTGTACTTCCCAA WTGCCACATTTTAAA		Other	3' UTR	
	HSTSCGENE	1392	C	0.97	T	0.03	0.06	ATCTTGGGGCCACYCTCTCTCTTCCCT	ACC	Synonymous	Thr	Thr
	HSTSCGENE	1881	G	0.98	A	0.02	0.03	GCCTCAGCTACTCRGTGGGCTCAATGA	TCG	Synonymous	Ser	Ser
	HSTSCGENE	2139	C	0.88	T	0.13	0.22	TCGGATGTCA TTCCYAGGAGCCTCCGCGAG	GCC	Synonymous	Ala	Ala
	HSTSCGENE	2595	C	0.90	T	0.10	0.18	CGTGTTGCTGAGGYGCCAGATTAAACAG	GCG	Synonymous	Gly	Gly
	HSTSCGENE	3269	G	0.81	A	0.19	0.30	GGTCITGTGTTTATRGCTAGAGAAATAG		Other	3' UTR	
	HSTSCGENE	3660	C	0.94	T	0.06	0.11	CTGCAACCTCTCCYGGTTC AAGCAITTT		Other	3' UTR	
	HSTSCGENE	3710	G	0.98	C	0.02	0.03	TAGCTGGGATTACASGCCCTGCCATCAC		Other	3' UTR	
	HSTSCGENE	3727	C	0.69	G	0.31	0.43	ACCTGCCATCACACASAGCTAATTTTGTGA		Other	3' UTR	
	HSTSCGENE	3838	G	0.90	A	0.10	0.18	CCCAAAGTCTGGGRTTACAGRCCTGAGC		Other	3' UTR	
	HUMAPNIIIA	3057	T	0.92	C	0.08	0.15	AGGGCATCTCTGAGYGTCTCTGCTGGAG		Other	3' UTR	
	HUMGFAT	2930	T	0.65	G	0.35	0.45	ATCTCTAAAAGTGTCTTTTATTTCTCTTG		Other	3' UTR	
	HUMGLUTRN	2110	G	0.94	C	0.06	0.12	GGCTATGGCCACCCCTTCTGCTGCCCTGG		Other	3' UTR	
	HUMGLUTRN	933	A	0.99	C	0.01	0.02	GATGATCGGGGAGAMGAGTCCATCC	AAG	Nonsynonymous	Lys	Thr
	HUMGUANCYC	2388	C	0.93	T	0.07	0.14	ATTGCACTGAATAYTGCTCTCGTGGAG	TAC	Synonymous	Tyr	Tyr
	HUMGUANCYC	2571	A	0.86	G	0.14	0.24	CGTTTGTGCTCAARATCACAGACTATGG	AAA	Nonsynonymous	Lys	Arg
	HUMGUANCYC	2643	G	0.93	A	0.07	0.14	GCCCTCTATGCCAARAAGCTGTGGACTGC	AAG	Synonymous	Lys	Lys
	HUMGUANCYC	2787	C	0.93	G	0.07	0.13	GAGGGCTGGACCTAGCCCCAAGAGAT	CTC	Synonymous	Leu	Leu
	IUMGUANCYC	2905	C	0.97	G	0.03	0.07	ACCGATGTTGGGCTAGGACCCAGCTGAG	CAG	Nonsynonymous	Gln	Glu
	IUMGUANCYC	3300	C	0.81	G	0.19	0.31	GACAACITTTGATGTSTACAGGTGGAGAC	GTC	Synonymous	Val	Val
	HUMGUANCYC	3663	G	0.96	A	0.04	0.08	CTTCGGGUGGATGTGROAAATGAAGGGA	GTC	Synonymous	Val	Val
	IAPPEX1-2	199	T	0.97	C	0.03	0.06	TTTATTTAGAGAAAYGCACACTTGTGTTT		Other	Intron	

NSDOCID: <EP_____0955382A2_I_>

55	KALSTEX3	91	T	0.97	C	0.03	0.07	CTGGCTCCTATGTATYTAGATCAGATTG	TTA	CTA	Synonymous	Leu	Leu
	KLKEX1	105	G	0.98	T	0.02	0.03	AGGCAATTCGAAGKCCAAGGCTTATAT			Other	5' UTR	5' UTR
	KLKEX3	253	C	0.68	G	0.32	0.44	TGAGTTGCCACCCAGGAACTTGAAGTG	CAG	GAG	Nonsynonymous	Gln	Gln
	KLKEX3	50	G	0.98	A	0.02	0.03	GCTCTGGCTGGTCTCCCAACTTGTTG	CGC	CAC	Nonsynonymous	Arg	Arg
	KLKEX4	110	T	0.97	A	0.03	0.07	CCACGTCAGAAAGGAGACAGACTTCATGC	GTG	GAG	Nonsynonymous	Val	Val
	KLKEX4	88	A	0.66	G	0.34	0.45	CTAATGATGAGTGCRAAAAAGCCACGTC	AAA	GAA	Nonsynonymous	Lys	Lys
	KLKEX5	318	A	0.75	G	0.25	0.38	CCCCAGCTGTGTCARTCTCATGGCCTGGA			Other	5' UTR	5' UTR
	MRLEX1B	156	T	0.46	C	0.54	0.50	CCGATCAGCCAAATAYTGGACTTGTGTTG			Other	5' UTR	5' UTR
	MRLEX1B	16	G	0.90	C	0.10	0.18	GGGCGGTGCCCGCSTCCCTCTGCGCG			Other	5' UTR	5' UTR
	MRLEX2	1338	A	0.97	C	0.03	0.05	CTCTTTTAAAGGGAMTCCACACGTAAACC	AAT	ACT	Nonsynonymous	Asn	Thr
	MRLEX2	1405	T	0.99	C	0.01	0.03	GATGATAAGACTAYTATTCCTATCAGG	TAT	TAC	Synonymous	Tyr	Tyr
	MRLEX2	1617	G	0.99	A	0.01	0.03	AATATCTTTATCACRATCGGCTAGAGACC	CGA	CAA	Nonsynonymous	Arg	Gln
	MRLEX2	1668	A	0.97	G	0.03	0.05	CTTCTCTCTGTCTARTACTTTAGTGGAGT	AAT	AGT	Nonsynonymous	Asn	Ser
	MRLEX2	1696	C	0.97	T	0.03	0.06	TCATGGAATACAYGGCGACCTGTCTGC	CAC	CAT	Synonymous	Ile	His
	MRLEX2	1720	T	0.50	C	0.50	0.50	TGCTCTAGAAAGAGYATGGGTATCCGGT	AGT	AGC	Synonymous	Ser	Ser
	MRLEX9	1326	C	0.99	T	0.01	0.03	GGAAATGACACACTGGYGGTGTCTGCAGCTC			Other	3' UTR	3' UTR
	MRLEX9	1572	A	0.43	G	0.57	0.49	GTTAAAGATCAGCTRTTCCCTTCTGATCT			Other	3' UTR	3' UTR
	MRLEX9	1670	A	0.88	G	0.13	0.22	GGCCCACTTGGCARGGTTTCAGTCTGAAT			Other	3' UTR	3' UTR
	MRLEX9	1964	G	0.86	A	0.14	0.24	AATCTTTTAAAAATRATGATAATCATCAG			Other	3' UTR	3' UTR
	MRLEX9	247	T	0.97	C	0.03	0.06	ACCTGTTTTTAAACAYGTGATGGTTGATTC			Other	3' UTR	3' UTR
	MRLEX9	2551	T	0.88	C	0.13	0.22	CCAAATGCTGTCTGCTCTTATTTTGT			Other	3' UTR	3' UTR
	MRLEX9	2635	G	0.36	A	0.64	0.46	TCATATAATTTAAARAACACTAAATTAG			Other	3' UTR	3' UTR
	MRLEX9	869	C	0.99	G	0.01	0.02	TTTCTGTCTGTASATTAATCTGTATGTAT			Other	3' UTR	3' UTR
	MRLEX9	916	A	0.83	C	0.18	0.29	AATAAGGTATAAGGMTCTTTTGTAAATGA			Other	3' UTR	3' UTR
	NCX1EX12	1135	A	0.57	G	0.43	0.49	AGATTCCAGGAACRTGCAAAATCCCTTC			Other	3' UTR	3' UTR
	NCX1EX12	1190	C	0.80	T	0.20	0.32	TGATTGGCAAGGTCYTTCTTCCAGCATTC			Other	3' UTR	3' UTR
	NCX1EX12	1298	A	0.99	G	0.01	0.03	ATAACCCCATTCARAAGCACATCATCGT			Other	3' UTR	3' UTR
	NCX1EX12	1366	G	0.99	C	0.01	0.03	CGTTGCTGGGATTTCTCTCTAGTTTTAT			Other	3' UTR	3' UTR
	NCX1EX12	1407	A	0.84	G	0.16	0.26	CCATGGCTTGCACARTCTGTTCCAGTCA			Other	3' UTR	3' UTR
	NCX1EX12	1841	G	0.97	C	0.03	0.06	ACCATTAATTTCAGSAAGGCCAAGGAGAA			Other	3' UTR	3' UTR
	NCX1EX12	2099	T	0.94	C	0.06	0.11	GAAAGAGCCAGGAGGACCAACGGGCTTT			Other	3' UTR	3' UTR

55	NCX1EX12	2123	T	0.70	Del	0.30	0.42	GCCTTTAAAGTGTGTCTCTCTCTACTTA	CGT	CAT	Nonsynonymous	Other	3' UTR
	NCX1EX12	2614	T	0.96	C	0.04	0.08	TGTGATTACTATTTTCATGAGTAAAGTG			Other	Other	3' UTR
	NCX1EX12	2810	G	0.67	A	0.33	0.44	TTTATCTTTGACCGCTTGCAGATAAATA			Other	Other	3' UTR
	NCX1EX12	2832	C	0.99	G	0.01	0.03	ATAATATATCTCTSCATTTTAAACCAAG			Other	Other	3' UTR
	NCX1EX12	3079	A	0.66	C	0.34	0.45	TAAACATTAGAAAAAMTTTTTGCACATTT			Other	Other	3' UTR
	NCX1EX12	3193	G	0.99	C	0.01	0.03	TTGAAAGCTTTTGTGTTTGTGCTTTT			Other	Other	3' UTR
	NCX1EX12	664	T	0.99	C	0.01	0.03	TCTCTCCAGGTGAYAAATCCTTAAAGCT			Other	Other	3' UTR
	NCX1EX12	709	A	0.94	G	0.06	0.10	TTGGTTTGTGTTTCRGTTGGAGCTGGGAG			Other	Other	3' UTR
	NCX1EX12	948	G	0.89	A	0.11	0.20	AGCATGCTTTCATCTATATTACCAAAATTC			Other	Other	3' UTR
	NCX1EX4	59	G	0.99	A	0.01	0.03	TAGAATATTTGACCCRTGAGGAATATGAGA			Other	Other	3' UTR
	NCX1EX9	66	A	0.97	T	0.03	0.06	ACTGACGACAAAGWGGAAAGAGAGAGGCG			Nonsynonymous	Other	Arg
	NETEX11	123	T	0.86	G	0.14	0.24	CGTCAGTCTGCTKCTCCTCGCTGGTGTGTA	TTC	TGC	Other	Intron	His
	NETEX12	81	T	0.93	C	0.07	0.13	TCACCTACGAGGACACATCTTCCCGCCC	TAC	CAC	Nonsynonymous	Phe	Cys
	NETEX13	50	G	0.93	A	0.07	0.12	GCCTATGGCATCACRCCAGAGAACGAGCA	ACG	ACA	Nonsynonymous	Tyr	His
	NETEX14	29	G	0.93	C	0.07	0.14	TGTCTTTCTCTGCASTTGGCAACTGGCT			Synonymous	Thr	Thr
	NETEX5	121	A	0.93	C	0.07	0.12	CTCCAATGGCATCMTGCCCTACTGTCACA	AAT	ACT	Other	Intron	Thr
	NETEX5	175	A	0.96	G	0.04	0.07	CACGGTCAGTCTCRGTGACCAACCAAGCC			Nonsynonymous	Asn	Thr
	NETEX5	83	C	0.95	G	0.05	0.10	TTCTGCTCTCTGGTSCATGGCTTCACGCT	GTC	GTC	Other	Intron	Thr
	NETEX7	112	G	0.92	C	0.08	0.15	TCCTTGGTTACATGSCCCATGAAACACAAAG	CCC	CCC	Synonymous	Val	Val
	NETEX7	131	A	0.93	G	0.07	0.14	TGAACACAAAGGTCAACATTTGAGGATGTGG	ACC	AGC	Nonsynonymous	Ala	Pro
	NETEX7	73	G	0.94	C	0.06	0.11	GTATCACCAGCTTCTCTCTGGGTTCCGC	GTC	CTC	Nonsynonymous	Asn	Ser
	NETEX8	17	C	0.55	A	0.45	0.49	TGATGAGGTCTTUMTGTTTCTTACAGGA			Nonsynonymous	Val	Leu
	NETEX9	157	A	0.91	G	0.09	0.16	GTCTGCATAACCAAGGTGAGTAGGGGCT	AAG	AGG	Other	Intron	Thr
	NETEX9	36	G	0.96	A	0.04	0.07	GAGCTGTCTACACRGGCTGGCAGATCA	ACG	ACA	Nonsynonymous	Lys	Arg
	NPYEX1	112	G	0.97	A	0.03	0.06	GCCTGGCCGAGGCRATACCCCTCCAAAGCC	GCG	GCA	Synonymous	Thr	Thr
	NPYEX1	178	A	0.90	G	0.10	0.18	GCCAGATATACTACRGGCTGGGACACTA	TCA	TCG	Synonymous	Ala	Ala
	NPYEX1	92	C	0.95	A	0.05	0.10	CCCTGCTGTGTGCMTGGTGGCTGGGCCC	CTG	ATG	Synonymous	Ser	Ser
	NPYEX2	45	T	0.40	C	0.60	0.48	TATGAAAAACGATCYAGCCACAGACACT	TCT	TCC	Nonsynonymous	Leu	Met
	NPYEX3	100	A	0.96	G	0.04	0.08	CCCTATTTCAGCCCTATTTCATCTGTGTA			Synonymous	Ser	Ser
	NPYEX3	78	G	0.91	T	0.09	0.16	GAGACTTGTCTCTKGGCTTTCCATATTT			Other	3' UTR	
	NPYEX3	144	T	0.94	G	0.06	0.12	AACATACTGTCCATKTGTCTAAAAATAATC			Other	3' UTR	
	NPYEX2										Other	5' UTR	

38

SDOCID: <EP_____0955382A2_1_>

40

	5	10	15	20	25	30	35	40	45	50	55			
PTGERJEX1	371	C	0.96	G	0.04	0.07	GC	CGCGGGGCA	ACTS	ACGCGCCCTCCAGG	CTC	CTG	Synonymous	Leu
PTGERJEX1	765	A	0.98	T	0.02	0.04	G	GTATGCGAGCC	ACWTG	AAGACGGCGTGC	ATG	TTG	Nonsynonymous	Leu
PTGERJEX1	878	G	0.90	T	0.10	0.18	C	CAGTGGCGGG	CGACKTGG	CTTCAACAG	ACG	ACT	Synonymous	Thr
PTGERJEX10	206	T	0.98	C	0.03	0.05	C	ACATGTTT	TGTAC	YTTTACTATATCTAC	.	.	Other	3' UTR
PTGERJEX10	281	A	0.85	G	0.15	0.26	G	CGGTATACAT	TATCTAT	GTAAAAATTTC	.	.	Other	3' UTR
PTGERJEX2	1293	T	0.38	C	0.62	0.47	C	ACTAAAAATG	TTTTTTC	TACAGTCTACATG	.	.	Other	Intron
PTGERJEX2	1295	T	0.86	C	0.14	0.23	C	TAAAAATG	TTTTTTC	YACAGTCTACATGAA	.	.	Other	Intron
PTGERJEX2	1393	T	0.84	C	0.16	0.27	C	GCACCTCT	TAAAAA	YGTCTCCGCCACCA	.	.	Other	Intron
PTGERJEX2	1403	C	0.98	A	0.03	0.05	A	AAATGTCT	CCCCCAM	CAACATAGTAATC	.	.	Other	Intron
PTGERJEX2	1614	T	0.94	C	0.06	0.11	C	TAAAGAAAT	TAATTT	YGATAGGTACAATAT	.	.	Other	Intron
PTGERJEX2	1719	G	0.98	C	0.03	0.05	C	TGGAGACAAA	ATCTT	STTGAGAGTGCTTAT	.	.	Other	Intron
PTGERJEX2	2153	A	0.99	G	0.01	0.03	A	AGTCCATCA	GCGCTORT	AAAGTGAATATTT	.	.	Other	Intron
PTGERJEX2	2517	T	0.92	C	0.08	0.15	C	TAGGCATTC	GTTAGYA	TGGGGAAACCTGA	.	.	Other	Intron
PTGERJEX2	3069	T	0.93	C	0.08	0.14	C	TAGTGTGTAT	ATAYCC	CAAGATATTTTA	.	.	Other	Intron
PTGERJEX2	3101	A	0.91	G	0.09	0.17	A	AAATGTAA	GTTTTRAT	CATGCCAGATTT	.	.	Other	Intron
PTGERJEX2	326	T	0.91	A	0.09	0.17	A	TATCGCTAA	ACCTWACT	GTGAATTTAGG	.	.	Other	Intron
PTGERJEX2	3282	A	0.98	G	0.03	0.05	A	CTAAAAA	CTGGCAR	CACGATTTTAAATA	.	.	Other	Intron
PTGERJEX2	3382	T	0.63	C	0.37	0.47	C	TTTTATAAT	TTTGTCT	TTTTGACTCCA	.	.	Other	Intron
PTGERJEX2	557	G	0.99	T	0.01	0.03	T	TATAAATGAT	CTTGKTC	TATGGGAGCG	.	.	Other	Intron
PTGERJEX2	628	T	0.83	C	0.17	0.28	C	AACCACAT	ACATCA	YTGAAAGACAAGGAT	.	.	Other	Intron
PTGERJEX2	769	T	0.91	A	0.09	0.17	A	GTATAATG	TATTA	WAATATTCATCGATA	.	.	Other	Intron
PTGERJEX2	787	T	0.94	G	0.06	0.12	G	ATTCATCG	ATACCA	KATTCAAAATATTC	.	.	Other	Intron
PTGERJEX2	805	A	0.91	C	0.09	0.16	C	TCAAAATAT	TGCTC	AMTACAGCAAAATTAGC	.	.	Other	Intron
PTGERJEX2	850	G	0.98	A	0.02	0.04	A	TTTAAGTT	ACTTGT	GTATTAATTAGGTTT	.	.	Other	Intron
PTGERJEX2	852	T	0.62	A	0.38	0.47	A	TAACTTTA	CTTGG	AWTGATAATTAGGTTT	.	.	Other	Intron
PTGERJEX2	855	A	0.98	T	0.02	0.04	T	TTTTACTT	GGATTG	WTAAATTAGGTTTACT	.	.	Other	Intron
PTGERJEX3	76	C	0.94	T	0.06	0.12	C	CTCCACCT	CTCTTAC	YCTGCCAGTTCCT	CCC	CTC	Nonsynonymous	Pro
PTGERJEX3	80	C	0.93	T	0.07	0.13	C	ACCTCTT	ACCTG	YCACTGTTCTCTCAAC	TGC	TGT	Synonymous	Cys
PTGERJEX4	719	G	0.84	T	0.16	0.27	T	TCTAAGCT	TTTGT	ATKACAAAGGATGATG	.	.	Other	3' UTR
PTGERJEX4	94	C	0.98	T	0.03	0.05	C	TTTGCATAT	TTCTT	YCCACCTGAGAAGGA	.	.	Other	3' UTR
PTGERJEX6	197	A	0.98	G	0.03	0.05	G	GAGTGTGT	GTGTTT	RAAAACCAAGCTCC	.	.	Other	3' UTR

55	PTGERJEX6	300	G	0.91	A	0.09	0.16	GAGATTACGACGAARCCAGGTCAATTTCCG	.	.	Other	3' UTR	.
	PTGERJEX6	387	T	0.98	A	0.03	0.05	CCAAATTAGACTTAAWAGTAAGAAATAGCAC	.	.	Other	3' UTR	.
	PTGERJEX7	85	A	0.98	G	0.02	0.04	TTGGTGCAGTTCCTCCTGATAGTGAAGTAC	CAT	CGT	Nonsynonymous	His	Arg
	PTGERJEX8	116	C	0.83	T	0.17	0.28	GATTTGCTCTTCCTCCGCAATGCTTCATC	CCC	CCT	Synonymous	Pro	Pro
	PTGERJEX9	16	T	0.94	C	0.06	0.12	TGCCTATCACATAAYAGGAGAACCTGCA	.	.	Other	Intron	.
	RENEX1	80	A		T			GGAACATGATGATGGWTGGAGAAAGATGCC	GGA	GGT	Synonymous	Gly	Gly
	RENEX2	135	A	0.76	C	0.24	0.37	ATGAAGAGGCTGACMCTTTGGCAACACCAC	ACA	ACC	Synonymous	Thr	Thr
	RENEX4	151	A	0.97	G	0.03	0.05	GACATCATCACCGTTRAGTTGGGCCGCCCT	.	.	Other	Intron	.
	RENEX4	165	T	0.66	G	0.34	0.45	AAGTTGGGCCGCCCKAGGTCACTGCCCCC	.	.	Other	Intron	.
	RENEX9	138	G	0.99	A	0.01	0.03	TTCAAGTGAGGTTCTCRAGTCGGCCCCCTCG	.	.	Other	Intron	.
	SAEX1	167	T	0.91	C	0.09	0.17	GTTTGGGCCAGTCYTGCTCCTCCGGATT	.	.	Other	Promoter	.
	SAEX1	76	G	0.98	T	0.03	0.05	ATTACCTGTAAAGAGKAACCGCTGGAGTC	.	.	Other	Promoter	.
	SAEX11	143	C	0.99	T	0.01	0.02	AGAGCAGATGATGYATATTAATCTCTGG	GTC	GTT	Synonymous	Val	Val
	SAEX2	54	T	0.94	C	0.06	0.12	CTCTGTGCAATCCYAGTGTCTAAAGCTT	.	.	Other	5' UTR	.
	SAEX3	109	T	0.99	C	0.01	0.03	GAGTTTGAAGAACYGGGATCTCTGTCCA	CTG	CCG	Nonsynonymous	Leu	Pro
	SAEX4	187	T	0.82	C	0.18	0.30	CACCTCAAAGCTGATYGTATCAGAGAACTC	ATT	ATC	Synonymous	Ile	Ile
	SAEX5	182	T	0.14	C	0.86	0.24	TGGAAGGTATACTTTCACAAAAGTGCAGC	.	.	Other	Intron	.
	SAEX8	111	G	0.97	C	0.03	0.06	AAATGGAGAAACAASACGGCTGGATAT	AAG	AAC	Nonsynonymous	Lys	Asn
	SAEX9	101	C	0.98	T	0.03	0.05	CCCTCTCTGCTTTTGTGATGTTAAGGTTTG	TTC	TTT	Synonymous	Phe	Phe
	SCNNIGEX1	167	G	0.48	A	0.52	0.50	GGTGCCCAAGGAAGRCGACGCGCGGCGG	.	.	Other	Promoter	.
	SCNNIGEX1	236	G	0.81	T	0.19	0.30	TGAAGTCGTGGCCCKCTCCGGCGGTCTC	.	.	Other	Promoter	.
	SCNNIGEX1	498	G	0.99	A	0.01	0.02	TGGACGGGATGCCGRCGCCAGGGCGTCG	.	.	Other	Intron	.
	SCNNIGEX1	552	C	0.70	G	0.30	0.42	GAGCCAGCATCAGCGGTGGLGGCTTCCC	.	.	Other	Intron	.
	SCNNIGEX1	553	G	0.99	A	0.01	0.02	AGCCAGCATCAGCCRGTTGGCGGCTTCCC	.	.	Other	Intron	.
	SCNNIGEX12	1016	T	0.80	A	0.20	0.32	AGATCAGAGTGCCGCGGTTGGAGGCTCGGG	.	.	Other	3' UTR	.
	SCNNIGEX12	1085	A	0.75	G	0.25	0.38	CAGGAGATGATTTTGTATTTCAAATTTTG	.	.	Other	3' UTR	.
	SCNNIGEX12	407	C	0.78	G	0.22	0.34	ATGCTGGATGAGCTSTGAGGACAGGTTGA	CTC	CTG	Synonymous	Leu	Leu
	SCNNIGEX12	454	G	0.99	T	0.01	0.02	GACCACAGCCATGKTCTAAGGACATGGA	.	.	Other	3' UTR	.
	SCNNIGEX12	485	G	0.99	A	0.01	0.02	GGTGCCCCCAGACRTGTGCACAGGGGAC	.	.	Other	3' UTR	.
	SCNNIGEX12	569	T	0.86	G	0.14	0.24	CGCAAGATGGGGCKGGGATGCGCAGGA	.	.	Other	3' UTR	.
	SCNNIGEX12	646	C	0.80	T	0.20	0.32	ATAAATCCCGGGACYTGAACCTATTAGCAC	.	.	Other	3' UTR	.

5DOCID: <EP_____0955382A2_1_>

5

10

15

20

25

30

35

40

45

50

55

TBXA2REX3	953	G	0.39	C	0.61	0.47	TTCAAGCGATTCTCTGCTCAGCCTCCC	GAC	GTC	Other	3' UTR	Val
TBXASEX10	61	G	0.97	T	0.03	0.06	CAGCTCGAGGAAGKCTGCCCTATCTGG	CTG	CTA	Nonsynonymous	Gly	Val
TBXASEX10	98	G	0.93	A	0.07	0.14	ATTGACAGACCGCTRAGGATGTACCCGCC	GCC	GCT	Synonymous	Leu	Leu
TBXASEX11	105	C	0.91	T	0.09	0.16	GTGCTTAGAGATGCGYGTGGTGGCCCTGCA	ACC	AAC	Synonymous	Ala	Ala
TBXASEX11	152	C	0.99	A	0.01	0.03	GCCAAGCCCGAGAMCTTCAACCCCTGAAA	CAG	GAG	Nonsynonymous	Thr	Asn
TBXASEX11	49	C	0.98	G	0.02	0.05	CACGGAGGCAGCTAGGACITGGAGGTTG	CGC	TGC	Synonymous	Gln	Glu
TBXASEX11	73	C	0.99	T	0.01	0.02	AGTGCTGGGGCAGYGGATCCCCCGAGGC	GCT	ACT	Nonsynonymous	Arg	Cys
TBXASEX11	88	G	0.90	A	0.10	0.18	GCATCCCCCGCAGGRCCTGTGCTAGAGATG	CGG	AGG	Nonsynonymous	Ala	Thr
TBXASEX12	46	C	0.98	A	0.02	0.04	TCACGGCTGAGGGCMGGCAGCAGCACCGG	CTT	CTC	Synonymous	Arg	Arg
TBXASEX13	226	A	0.99	G	0.01	0.02	CCTGGCATGCCAAGGRTAAGAGGTTCTTTT	CTT	CTT	Other	3' UTR	
TBXASEX4	130	C	0.99	T	0.01	0.03	CCAACAGAAATGGTAYGTAGTTTCTTTTC	CTT	CTT	Other	Intron	
TBXASEX5	15	G	0.99	A	0.01	0.02	CTGACCCCTCTGCTTCTTACTTCCCAACAG	CTT	CTT	Other	Intron	
TBXASEX6	59	C	0.98	A	0.02	0.04	ACCCAAGCCTGGAMCTTCTCTGGCTCA	GAC	GAA	Synonymous	Asp	Asp
TBXASEX8	110	A	0.89	G	0.11	0.20	ATGGCTTTTAAACRAACTCATTAGGAAT	AAA	GAA	Nonsynonymous	Lys	Asp
TBXASEX8	119	A	0.99	G	0.01	0.03	TTAAACAAACTCATTRGGAAATGTGATGGC	AGG	GGG	Nonsynonymous	Arg	Gly
TBXASEX9	156	C	0.96	A	0.04	0.07	CGAACCCCTTCCCGMAACACACAGCCACGC	CAA	AAA	Nonsynonymous	Gln	Lys
TBXASEX9	276	C	0.88	G	0.12	0.21	CTTTGGCCACTACTACTGGCCACCAAC	CTA	GTA	Nonsynonymous	Leu	Val
TRIUREX1	56	G	0.44	C	0.56	0.49	TTCTGCAGAACTTASATGATAAGCAACGA	CTA	CTA	Other	Promoter	
TRIUREX1	84	T	0.98	C	0.02	0.04	ACAAAGCCAGCTGCTCTAGACCCCTGGC	CTA	CTA	Other	Promoter	
TRIUREX2	147	C	0.99	A	0.01	0.03	GTCAGTGAACCTGAAMCAACACACAGCTTCA	AAC	AAA	Nonsynonymous	Asn	Lys
TRIUREX2	240	A	0.99	G	0.01	0.03	GGCCTGGGCAATGTGGCAACATCATGCT	GTA	GTC	Synonymous	Val	Val
TRIUREX3	1161	T	0.58	C	0.42	0.49	TCCCACATGATGGGGYGGAAAAAGGCCAAAA	CTA	CTA	Other	3' UTR	
TRIUREX3	1231	T	0.98	C	0.03	0.05	TTAAATTGAAAAGYATAGTCAAGACAAA	CTA	CTA	Other	3' UTR	
TRIUREX3	1540	T	0.97	A	0.03	0.06	TTCTTTTGTGTTTWTCTCAAAATGCTAGT	CTA	CTA	Other	3' UTR	
TRIUREX3	1786	A	0.99	T	0.01	0.03	GAATCTCCGAGGGCWAAAAATGCCCTTGG	CTA	CTA	Other	3' UTR	
TRIUREX3	1846	T	0.94	C	0.06	0.12	GTAGATCAAAAAAGYACCCATACCTTTAC	CTA	CTA	Other	3' UTR	
TRIUREX3	2046	G	0.98	A	0.02	0.04	CCTCATCTAGAGTTCGGCTTTTITTTTTT	CTA	CTA	Other	3' UTR	
TRIUREX3	2175	A	0.97	G	0.03	0.06	ACCTGCATGACAGTACAGCAATCTATGTTA	CTA	CTA	Other	3' UTR	
TRIUREX3	2283	G	0.95	A	0.05	0.10	ACAACCATGTGTRTTTATAAACACATA	CTA	CTA	Other	3' UTR	
TRIUREX3	377	T	0.95	C	0.05	0.10	GCCACAAAAGTGTCTTTGATGACACCTTG	CTT	TCC	Synonymous	Ser	Ser
TRIUREX3	960	T	0.96	C	0.04	0.07	TAAGATTTIAGACAYACATGTTAACTGTA	CTT	CTT	Other	3' UTR	

TABLE 2

Gene/ExOn	Base Position	Ref		Freq		Hetero- zygosity (H)	Sequence Tag		Ref Codon	Alt Codon	Type of amino acid change	Ref amino acid	Alt amino acid
		Allele	Allele	(P)	(Q)								
ACEEX13	138	C	T	0.81	0.19	0.30	CCTCTGCTGTCCTCCYAGCCAGGAGGCATC		CCC	CCT	Synonymous	Pro	Pro
ACEEX17	52	A	G	0.20	0.80	0.32	AATGTGATGGCCACRGTCCCGAAATATGA		ACA	ACG	Synonymous	Thr	Thr
ADRDJEX1	416	T	C	0.90	0.10	0.18	TGCTGGCCATCGCCYGGACTCCGAGACTC		TGG	CGC	Nonsynonymous	Trp	Arg
AGTEX2	644	C	T	0.86	0.14	0.24	UCTGCTGCTGTCCAYGGTGTGGGGTGT		ACG	ATG	Nonsynonymous	Thr	Met
AGTEX2	827	T	C	0.10	0.90	0.18	TGGCTGCTCCCTGAYGGGAGCCAGTGTGG		ATG	ACT	Nonsynonymous	Met	Thr
AGTEXP1	173	C	T	0.71	0.29	0.41	TGCTTGCTGTGTTTTCAGTGTCTATTA				Other	Promoter	
AGTEXP2	203	G	A	0.86	0.14	0.24	CTCGACCTTGCACACRGTCTCACTGTGTCA				Other	Promoter	
AGTEXP3	144	C	A	0.24	0.76	0.37	GCTATAAATAGGGCMTCGTGACCCGCCA				Other	Promoter	
ANPEX3	120	T	C	0.91	0.09	0.16	GTCTCTGCTGCAITTYGTGTCTCTTTGTTG				Other	3'UTR	
ANPEX3	33	T	C	0.80	0.20	0.32	TCTCTTGCAGTACYGAGATAACAGCCA		TGA	AGA	Nonsynonymous	Stop	Arg
ATIEX5	1138	A	G	0.93	0.07	0.13	AAGAAGCCTGCACCRGTGTTTGAGGTTGA		CCA	CCG	Synonymous	Pro	Pro
ATIEX5	1593	G	T	0.88	0.12	0.21	AAAGTTTCGTGCKGTTTTCAGCTATTA				Other	3'UTR	
ATIEX5	649	T	C	0.61	0.39	0.47	CAAAATTCAACCTYCCGATAGGCTGGG		CTT	CTC	Synonymous	Leu	Leu
MRLEX2	1504	C	T	0.89	0.11	0.20	CAAGAACAGATGAYGGGAGCTATTACCC		GAC	GAT	Synonymous	Asp	Asp
MRLEX2	545	A	G	0.81	0.19	0.30	CGCTCATGCGCCCRITGTTTAAAGCCCT		ATT	GTT	Nonsynonymous	Ile	Val
NCXIE12	3101	A	T	0.16	0.84	0.26	ACTCATTTTTAUCWGTATTAGGAATGTC				Other	3'UTR	

Table 3

Gene/Exon	Gene Name
Table 1	
AADD	Alpha-Adducin
ACE	Angiotensin Converting Enzyme
ADDB	Beta Adducin
ADDG	Gamma Adducin
ADORA2A	A2a Adenosine Receptor
ADRB3	Beta-3-Adrenergic Receptor
ADROM	(prepro)Adrenomedullin
AE1	Anion Exchanger
AGT	Angiotensinogen
ALDRED	Aldose Reductase
ANPEX1	Atrial Natriuretic Factor
APOA1	Apolipoprotein A-I
APOA2	Apolipoprotein A-II
APOA4	Apolipoprotein A-IV
APOC1EX1	Apolipoprotein C-I
APOC2	Apolipoprotein C-II
APOC3	Apolipoprotein C-III
APOC4	Apolipoprotein C-IV
APOER2	Apolipoprotein E Receptor 2
AT1	Angiotensin II Receptor Type-I
AT2	Angiotensin II Receptor Type 2
AVP	Arginine Vasopressin
AVPR2	Arginine Vasopressin Receptor Type II
BIR	Beta Inward Rectifier Subunit (Pancreatic K Channel)
BKRB2	B2-Bradykinin Receptor
BNP	Brain Natriuretic Protein

BRS3	Bombesin Receptor Subtype-3
CAL/CGRP	Calcitonin/Calcitonin Gene Related Peptide
C11Y	Chymase
CLCNKB	Chloride Channel (Human Kidney - B)
CNP	C-Type Natriuretic Peptide
COX1	Cytochrome P-450 11 Beta 1
COX2	Cytochrome P-450 11 Beta 2
CYP11B1	Dopamine Beta-Hydroxylase
CYP11B2	Dopamine D1 Receptor
DBH	Endothelin Receptor Subtype A
DDIR	Endothelin Receptor Subtype B
EDNRA	Endothelial Leukocyte Adhesion Molecule 1
EDNRB	Endothelin-2
ELAM1	Endothelin-1
ENDOTHEL	Galanin Receptor
ET1	Glucagon Receptor
GALNR	Growth Hormone 1
GGR	Growth Hormone 2
GHI	Glucose Insulinotropic Peptide Receptor or Gastric Inhibitory Polypeptide Receptor
GH2	Glucose Transporter 2
GIPR	Glucose Transporter 4
GLUT2	Glucose Transport-Like 5
GLUT4	G-Protein Beta-3 Chain
GLUT5	Glycogen Synthetase
GNB3	Haptoglobin
GSY1	Hydroxysteroid Dehydrogenase 11 Beta Kidney Isozyme
HAPT	Homo sapiens Thiazide-Sensitive Cotransporter
HSD11K	
HSTSCGENE	

HUMAPNIIIA	Human Na ⁺ /H ⁺ Antiporter
HUMGFAT	Human Glutamine:Fructose-6-Phosphate Amidotransferase
HUMGLTRN	Human Glucose Transporter
HUMGUANCYC	Human Guanylate Cyclase
IAPP	Islet Amyloid Polypeptide
ICAM1	Intercellular Adhesion Molecule 1
ICAM2	Intercellular Adhesion Molecule 2
INS	Insulin
KALST	Kallistatin
KLK	Kallikrein
MRL	Mineralocorticoid Receptor
NCX1	Sodium-Calcium Exchanger
NET	Norepinephrine Transporter
NPY	Neuropeptide Y
NPYR1	Neuropeptide Y Y1 Receptor
PGIS	Prostacyclin Synthase
PLA2A	Pancreatic Phospholipase A-2
PNMT	Phenylethanolamine N-Methyltransferase
PPGLUC	Preproglucagon
PPTHIR	Preprothyrotropin-Releasing Hormone
PTGER3	Prostaglandin E Receptor EP3 Subtype
REN	Renin
SA	SA Gene Acetyl-CoA Synthetase Homologue ?? (a candidate gene for genetic hypertension)
SCNNIG	Amloride-Sensitive Epithelial Sodium Channel Gamma Subunit
TBXA2R	Thromboxane A2 Receptor
TBXAS0	Thromboxane Synthase
TRHIR	Thyrotropin-Releasing Hormone Receptor
Table 2	
ACE	Angiotensin Converting Enzyme

ADRB3	Beta-3-Adrenergic Receptor
AGT	Angiotensinogen
ANP	Atrial Natriuretic Factor
ATI	Angiotensin II Receptor Type-I
MRL	Mineralocorticoid Receptor
NCX1	Sodium-Calcium Exchanger

Claims

1. A nucleic acid of between 10 and 100 bases comprising at least 10 contiguous nucleotides including a polymorphic site from a sequence shown in Table 1, column 8 or the complement thereof.
2. The nucleic acid of claim 1 that is DNA.
3. The nucleic acid of claim 1 that is RNA.
4. The nucleic acid of claim 1 that is less than 50 bases.

5. The nucleic acid of claim 1 that is less than 20 bases.
6. The nucleic acid of claim 1, wherein the polymorphic form occupying the polymorphic site is a reference base shown in Table 1, column 3.
7. The nucleic acid of claim 1, wherein the polymorphic form occupying the polymorphic site is an alternative base shown in Table 1, column 5.
8. The nucleic acid of claim 7, wherein the alternative base correlates with hypertension or susceptibility thereto.
9. The nucleic acid of claim 1, wherein the polymorphic site is one for which reference and alternative bases shown in columns 3 and 5 of Table 1 are respectively components of different codons encoding different amino acids.
10. The nucleic acid of claim 1, which is from a gene encoding a human angiotensin I receptor.
11. The nucleic acid of claim 1, which is from a gene encoding an angiotensin II receptor.
12. The nucleic acid of claim 1, which is from a gene encoding an atrial natriuretic peptide.
13. The nucleic acid of claim 1, which is from a gene encoding a β -3-adrenergic receptor.
14. The nucleic acid of claim 1, which is from a gene encoding a bradykinin receptor B2.
15. The nucleic acid of claim 1, which is from a gene encoding a mineralocorticoid receptor.
16. The nucleic acid of claim 1, which is from a gene encoding a renin protein.
17. The nucleic acid of claim 1, which from a gene encoding an angiotensinogen protein.
18. The nucleic acid of claim 1, which from a gene encoding a sodium calcium ion channel.
19. The nucleic acid of claim 1, which is from a gene encoding an angiotensin converting protein.
20. The nucleic acid of claim 1, which is from a gene encoding an angiotensin converting protein.
21. Allele-specific oligonucleotide that hybridizes to a sequence including a polymorphic site shown in Table 1 or the complement thereof.
22. The allele-specific oligonucleotide of claim 21 that is a probe.
23. An isolated nucleic acid comprising a sequence of Table 1, column 8 or the complement thereof, wherein the polymorphic site within the sequence or its complement is occupied by a base other than the reference base shown in Table 1, column 3.
24. A method of analyzing a nucleic acid, comprising:
 - obtaining the nucleic acid from an individual; and
 - determining a base occupying any one of the polymorphic sites shown in Table 1 or other polymorphic sites in equilibrium dislinkage therewith.
25. The method of claim 24, wherein the determining comprises determining a set of bases occupying a set of the polymorphic sites shown in Table 1.
26. The method of claim 25, wherein the nucleic acid is obtained from a plurality of individuals, and a base occupying one of the polymorphic positions is determined in each of the individuals, and the method further comprising testing each individual for the presence of a disease phenotype, and correlating the presence of the disease phenotype with the base.

27. The method of claim 24, wherein the determined base is correlated with susceptibility to hypertension.

28. A method of diagnosing a phenotype comprising:

determining which polymorphic form(s) are present in a sample from a subject at one or more polymorphic sites shown in Table 1;
diagnosing the presence of a phenotype correlated with the form(s) in the subject.

29. The method of claim 28, wherein the phenotype is hypertension.

30. A method of screening for a polymorphic site suitable for diagnosing a phenotype, comprising:

identifying a polymorphic site linked to a polymorphic site shown in Table 1, wherein a polymorphic form of the polymorphic site shown in Table 1 has been correlated with a phenotype; and
determining haplotypes in a population of individuals to indicate whether the linked polymorphic site has a polymorphic form in equilibrium dislinkage with the polymorphic form correlated with the phenotype.

31. The method of claim 30, wherein the polymorphic form of the polymorphic site shown in Table 1 has been correlated with hypertension.

32. The method of claim 30, wherein the linked polymorphic site and the polymorphic site shown in Table 1 are from the same gene.

33. A computer-readable storage medium for storing data for access by an application program being executed on a data processing system, comprising:

a data structure stored in the computer-readable storage medium, the data structure including information resident in a database used by the application program and including:
a plurality of records, each record of the plurality comprising information identifying a polymorphisms shown in Table 1.

34. The computer-readable storage medium of claim 33, wherein each record has a field identifying a base occupying a polymorphic site and a location of the polymorphic site.

35. The computer-readable storage medium of claim 33, wherein each record identifies a nucleic acid segment of between 10 and 100 bases from a fragment shown in Table 1 including a polymorphic site, or the complement of the segment.

36. The computer-readable storage medium of claim 33, comprising at least 10 records, each record comprising information identifying a different polymorphism shown in Table 1.

37. The computer-readable storage medium of claim 33, comprising at least 10 records, each record comprising information identifying a different polymorphism shown in Table 1.

38. A signal carrying data for access by an application program being executed on a data processing system, comprising:

a data structure encoded in the signal, said data structure including information resident in a database used by the application program and including:
a plurality of records, each record of the plurality comprising information identifying a polymorphism shown in Table 1

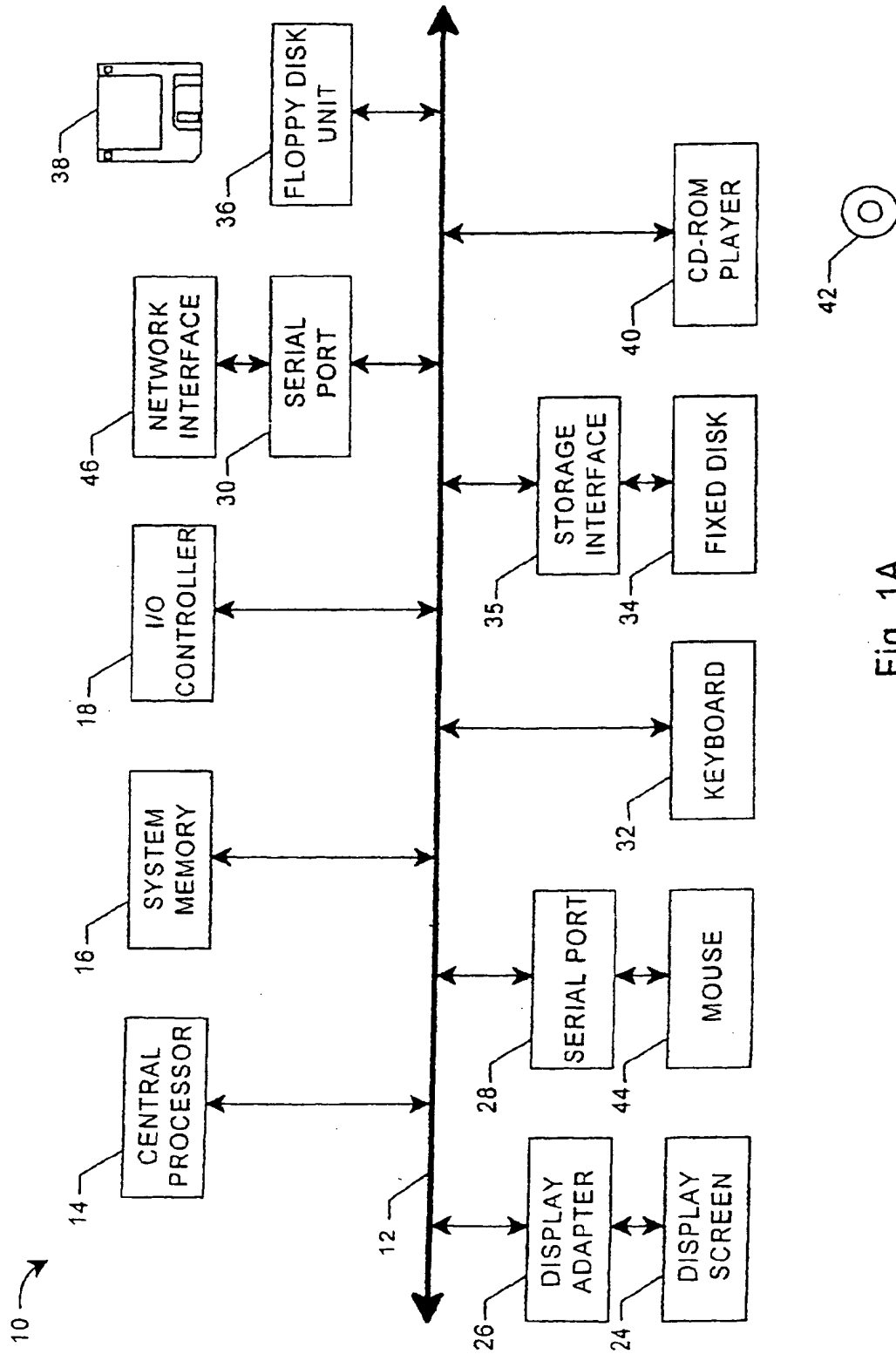


Fig. 1A

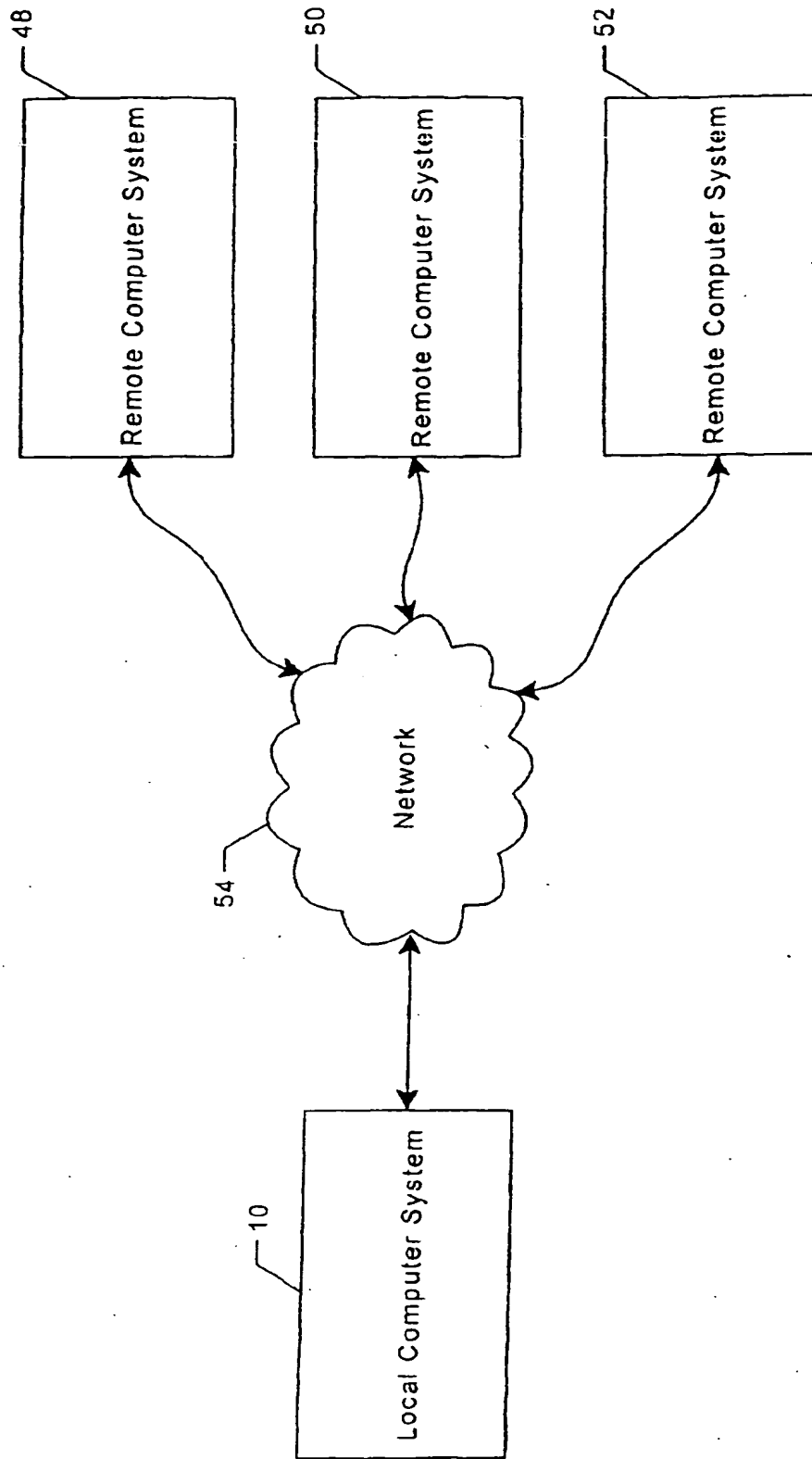
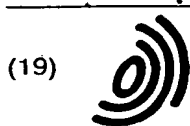


Fig. 1B





Europäisches Patentamt
European Patent Office
Office européen des brevets



(11)

EP 0 955 382 A3

(12)

EUROPEAN PATENT APPLICATION

(88) Date of publication A3:
05.07.2000 Bulletin 2000/27

(51) Int Cl.7: C12Q 1/68, C07K 14/72,
G06F 17/30

(43) Date of publication A2:
10.11.1999 Bulletin 1999/45

(21) Application number: 99250150.2

(22) Date of filing: 07.05.1999

(84) Designated Contracting States:
AT BE CH CY DE DK ES FI FR GB GR IE IT LI LU
MC NL PT SE
Designated Extension States:
AL LT LV MK RO SI

(30) Priority: 07.05.1998 US 84641 P
03.05.1999 US 304232

(71) Applicants:
• Affymetrix, Inc. (a California Corporation)
Santa Clara, CA 95051 (US)
• Case Western Reserve University
Cleveland, Ohio 44106 (US)

(72) Inventors:

- Fan, Jian Bing
San Diego, CA 92130 (US)
- Chakravarti, Aravinda
Shaker Heights, Ohio 44120 (US)
- Halushka, Mark Kenneth
Cleveland Heights, Ohio 44106 (US)

(74) Representative: UEXKÜLL & STOLBERG,
Patentanwälte
Beselerstrasse 4
22607 Hamburg (DE)

(54) Polymorphisms associated with hypertension

(57) The invention discloses a collection of polymorphic sites in genes known or suspected to have a role in hypertension. The invention provides nucleic acids including such polymorphic sites. The nucleic acids can

be used as probes or primers or for expressing variant proteins. The invention also provides methods of analyzing the polymorphic forms occupying the polymorphic sites.

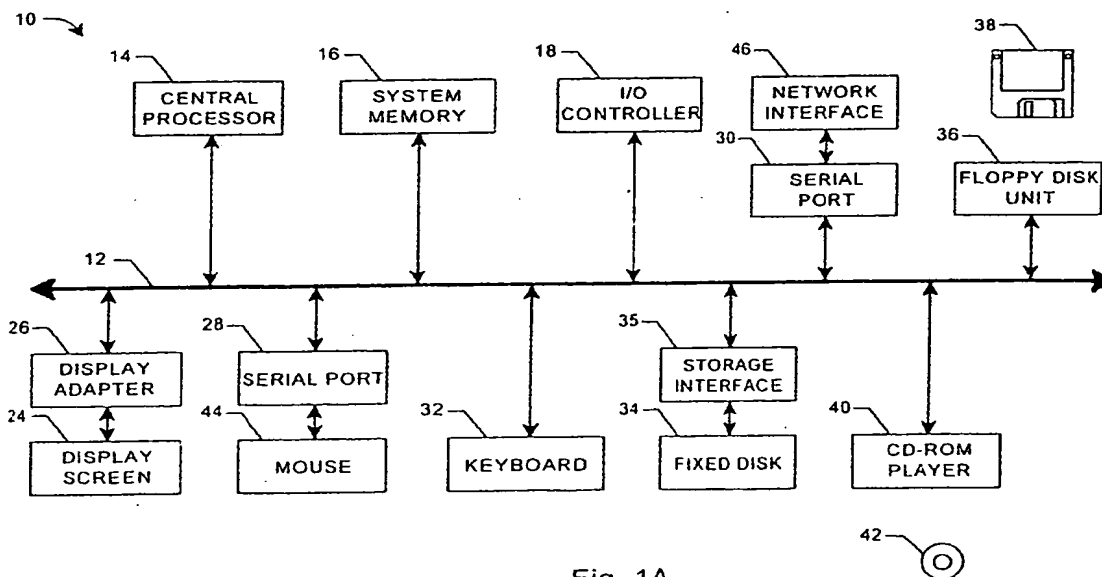


Fig. 1A

EP 0 955 382 A3

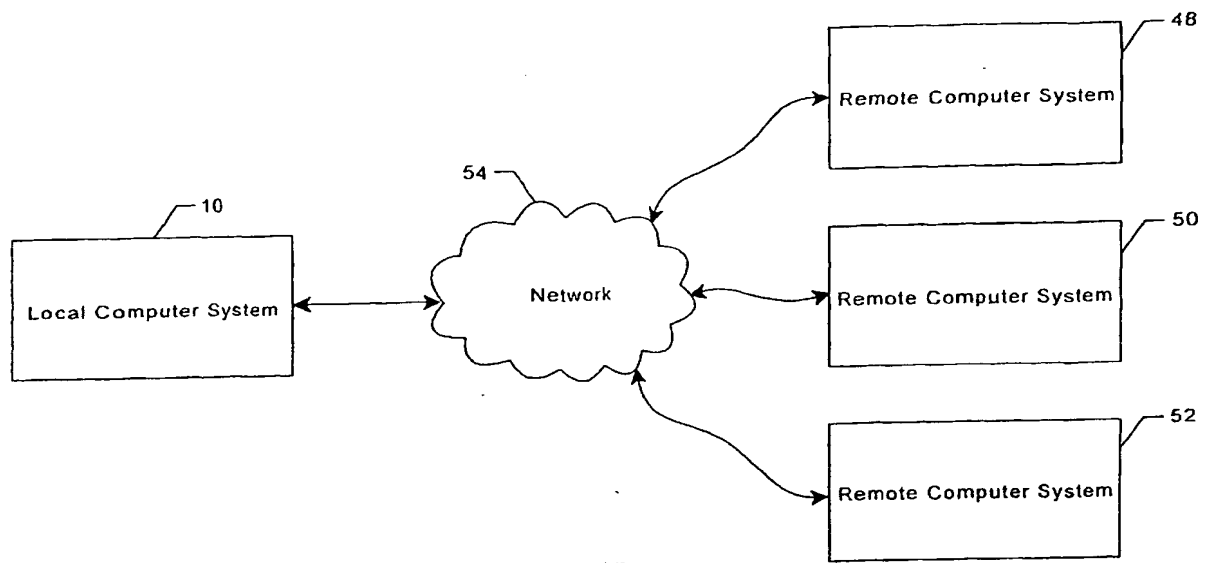


Fig. 1B



European Patent
Office

EUROPEAN SEARCH REPORT

Application Number
EP 99 25 0150

DOCUMENTS CONSIDERED TO BE RELEVANT			
Category	Citation of document with indication, where appropriate, of relevant passages	Relevant to claim	CLASSIFICATION OF THE APPLICATION (Int.CI.6)
X	WO 94 21790 A (UNIV BRITISH COLUMBIA) 29 September 1994 (1994-09-29) * see especially primer sequences ADU1524 and ADU133 within table 2 * * abstract; table 2 *	1,2,4,6,9	C12Q1/68 C07K14/72 G06F17/30
Y	CUSI D ET AL.: "Polymorphisms of alpha-adducin and salt sensitivity in patients with essential hypertension" THE LANCET, vol. 349, 1997, pages 1353-1357, XP002127732 * the whole document *	21-32	
D,Y	JEUNEMAITRE X ET AL.: "Haplotypes of angiotensinogen in essential hypertension" AMERICAN JOURNAL OF HUMAN GENETICS, vol. 60, 1997, pages 1448-1460, XP000857377 * abstract *	21-32	
Y	WO 94 08048 A (INST NAT SANTE RECH MED ;UNIV UTAH RES FOUND (US)) 14 April 1994 (1994-04-14) * the whole document *	21-32	TECHNICAL FIELDS SEARCHED (Int.CI.6) C12Q C07K
A	TAMAKI S ET AL.: "Polymorphism of alpha-adducin in japanese patients with essential hypertension" HYPERTENSION RESEARCH, vol. 21, 1998, pages 29-32, XP000862816 * the whole document *		
-/--			
The present search report has been drawn up for all claims			
Place of search THE HAGUE		Date of completion of the search 14 January 2000	Examiner Knehr, M
<p>CATEGORY OF CITED DOCUMENTS</p> <p>X : particularly relevant if taken alone Y : particularly relevant if combined with another document of the same category A : technological background O : non-written disclosure P : intermediate document</p> <p>T : theory or principle underlying the invention E : earlier patent document, but published on, or after the filing date D : document cited in the application L : document cited for other reasons & : member of the same patent family, corresponding document</p>			

EPO FORM 1503 03.82 (P04C01)



European Patent
Office

Application Number

EP 99 25 0150

CLAIMS INCURRING FEES

The present European patent application comprised at the time of filing more than ten claims.

- ☐ Only part of the claims have been paid within the prescribed time limit. The present European search report has been drawn up for the first ten claims and for those claims for which claims fees have been paid, namely claim(s):
- ☐ No claims fees have been paid within the prescribed time limit. The present European search report has been drawn up for the first ten claims.

LACK OF UNITY OF INVENTION

The Search Division considers that the present European patent application does not comply with the requirements of unity of invention and relates to several inventions or groups of inventions, namely:

see sheet B

- ☐ All further search fees have been paid within the fixed time limit. The present European search report has been drawn up for all claims.
- ☐ As all searchable claims could be searched without effort justifying an additional fee, the Search Division did not invite payment of any additional fee.
- ☐ Only part of the further search fees have been paid within the fixed time limit. The present European search report has been drawn up for those parts of the European patent application which relate to the inventions in respect of which search fees have been paid, namely claims:
- ☒ None of the further search fees have been paid within the fixed time limit. The present European search report has been drawn up for those parts of the European patent application which relate to the invention first mentioned in the claims, namely claims:

1-9, 21-38 (partially), see invention 1.



European Patent
Office

EUROPEAN SEARCH REPORT

Application Number
EP 99 25 0150

DOCUMENTS CONSIDERED TO BE RELEVANT			
Category	Citation of document with indication, where appropriate, of relevant passages	Relevant to claim	CLASSIFICATION OF THE APPLICATION (InCL6)
A	BIANCHI G ET AL.: "Two point mutations within the adducin genes are involved in blood pressure variation" PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCES USA, vol. 91, 1994, pages 3999-4003, XP002127733 * see especially Fig. 1 * * the whole document *		
A	WO 88 08457 A (BIOTECH RES PARTNERS LTD) 3 November 1988 (1988-11-03) * the whole document *		
D,A	LIFTON R P: "Genetic determinants of human hypertension" PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCES USA, vol. 92, 1995, pages 8545-8551, XP002127734 * the whole document *		TECHNICAL FIELDS SEARCHED (InCL6)
T	HALUSKA M K ET AL.: "Patterns of single-nucleotide polymorphisms in candidate genes for blood-pressure homeostasis" NATURE GENETICS, vol. 22, 1999, pages 239-247, XP002127735 * the whole document *	1-9, 21-38	
-The present search report has been drawn up for all claims			
Place of search THE HAGUE		Date of completion of the search 14 January 2000	Examiner Knehr, M
CATEGORY OF CITED DOCUMENTS X : particularly relevant if taken alone Y : particularly relevant if combined with another document of the same category A : technological background O : non-written disclosure P : intermediate document		T : theory or principle underlying the invention E : earlier patent document, but published on, or after the filing date D : document cited in the application L : document cited for other reasons & : member of the same patent family, corresponding document	

EPO FORM 1503 03 82 (P04C01)



European Patent
Office

LACK OF UNITY OF INVENTION
SHEET B

Application Number
EP 99 25 0150

The Search Division considers that the present European patent application does not comply with the requirements of unity of invention and relates to several inventions or groups of inventions, namely:

1. Claims: 1-9,21-38 (partial)

INVENTION 1:

A nucleic acid from the gene encoding alpha-adducin including a polymorphic site according to table 1, column 8, or the complement thereof, an allele-specific oligonucleotide hybridizing to such a polymorphic site, a method of analyzing such a nucleic acid, methods of diagnosing a phenotype related to such a polymorphic site, a computer-readable storage medium as well as a signal carrying data.

2. Claims: 1-9,21-38 (partial); 19,20 (complete)

INVENTION 2:

A nucleic acid from the gene encoding angiotensin converting enzyme including a polymorphic site according to table 1, column 8, or the complement thereof, an allele-specific oligonucleotide hybridizing to such a polymorphic site, a method of analyzing such a nucleic acid, methods of diagnosing a phenotype related to such a polymorphic site, a computer-readable storage medium as well as a signal carrying data.

3. Claims: 1-9,21-38 (partial)

INVENTION 3 TO INVENTION 5:

A nucleic acid from the genes encoding beta-adducin, gamma-adducin, or A2a adenosine receptor, including a polymorphic site according to table 1, column 8, or the complement thereof, an allele-specific oligonucleotide hybridizing to such a polymorphic site, a method of analyzing such a nucleic acid, methods of diagnosing a phenotype related to such a polymorphic site, a computer-readable storage medium as well as a signal carrying data.

Invention 3 refers to beta-adducin,
invention 4 refers to gamma-adducin, and
invention 5 refers to A2a adenosine receptor.

4. Claims: 1-9,21-38 (partial); 13 (complete)

INVENTION 6:

A nucleic acid from the gene encoding beta-3-adrenergic receptor including a polymorphic site according to table 1, column 8, or the complement thereof, an allele-specific oligonucleotide hybridizing to such a polymorphic site, a



European Patent
Office

LACK OF UNITY OF INVENTION
SHEET B

Application Number
EP 99 25 0150

The Search Division considers that the present European patent application does not comply with the requirements of unity of invention and relates to several inventions or groups of inventions, namely:

method of analyzing such a nucleic acid, methods of diagnosing a phenotype related to such a polymorphic site, a computer-readable storage medium as well as a signal carrying data.

5. Claims: 1-9,21-38 (partial)

INVENTION 7 TO INVENTION 8:

A nucleic acid from the genes encoding (prepro)adrenomedullin, or anion exchanger, including a polymorphic site according to table 1, column 8, or the complement thereof, an allele-specific oligonucleotide hybridizing to such a polymorphic site, a method of analyzing such a nucleic acid, methods of diagnosing a phenotype related to such a polymorphic site, a computer-readable storage medium as well as a signal carrying data.

Invention 7 refers to (prepro)adrenomedullin, and invention 8 refers to anion exchanger.

6. Claims: 1-9,21-38 (partial); 17 (complete)

INVENTION 9:

A nucleic acid from the gene encoding angiotensinogen including a polymorphic site according to table 1, column 8, or the complement thereof, an allele-specific oligonucleotide hybridizing to such a polymorphic site, a method of analyzing such a nucleic acid, methods of diagnosing a phenotype related to such a polymorphic site, a computer-readable storage medium as well as a signal carrying data.

7. Claims: 1-9,21-38 (partial)

INVENTION 10:

A nucleic acid from the gene encoding aldose reductase including a polymorphic site according to table 1, column 8, or the complement thereof, an allele-specific oligonucleotide hybridizing to such a polymorphic site, a method of analyzing such a nucleic acid, methods of diagnosing a phenotype related to such a polymorphic site, a computer-readable storage medium as well as a signal carrying data.

8. Claims: 1-9,21-38 (partial); 12 (complete)

INVENTION 11:



European Patent
Office

LACK OF UNITY OF INVENTION
SHEET B

Application Number

EP 99 25 0150

The Search Division considers that the present European patent application does not comply with the requirements of unity of invention and relates to several inventions or groups of inventions, namely:

A nucleic acid from the gene encoding atrial natriuretic factor including a polymorphic site according to table 1, column 8, or the complement thereof, an allele-specific oligonucleotide hybridizing to such a polymorphic site, a method of analyzing such a nucleic acid, methods of diagnosing a phenotype related to such a polymorphic site, a computer-readable storage medium as well as a signal carrying data.

9. Claims: 1-9,21-38 (partial)

INVENTION 12 TO INVENTION 19:

A nucleic acid from the genes encoding apolipoprotein A-I, apolipoprotein A-II, ..., apolipoprotein C-IV, or apolipoprotein E receptor 2, including a polymorphic site according to table 1, column 8, or the complement thereof, an allele-specific oligonucleotide hybridizing to such a polymorphic site, a method of analyzing such a nucleic acid, methods of diagnosing a phenotype related to such a polymorphic site, a computer-readable storage medium as well as a signal carrying data.

Invention 12 refers to apolipoprotein A-I,
invention 13 refers to apolipoprotein A-II,

...
invention 18 refers to apolipoprotein C-IV, and
invention 19 refers to apolipoprotein E receptor 2.

10. Claims: 1-9,21-38 (partial); 11 (complete)

INVENTION 20:

A nucleic acid from the genes encoding angiotensin II receptor type-1 or angiotensin II receptor type-2 including a polymorphic site according to table 1, column 8, or the complement thereof, an allele-specific oligonucleotide hybridizing to such a polymorphic site, a method of analyzing such a nucleic acid, methods of diagnosing a phenotype related to such a polymorphic site, a computer-readable storage medium as well as a signal carrying data.

11. Claims: 1-9,21-38 (partial)

INVENTION 21 TO INVENTION 23:

A nucleic acid from the genes encoding arginine vasopressin, arginine vasopressin receptor type II, or beta inward rectifier subunit (pancreatic K channel), including a polymorphic site according to table 1, column 8, or the complement thereof, an allele-specific oligonucleotide



European Patent
Office

LACK OF UNITY OF INVENTION
SHEET B

Application Number
EP 99 25 0150

The Search Division considers that the present European patent application does not comply with the requirements of unity of invention and relates to several inventions or groups of inventions, namely:

hybridizing to such a polymorphic site, a method of analyzing such a nucleic acid, methods of diagnosing a phenotype related to such a polymorphic site, a computer-readable storage medium as well as a signal carrying data.

Invention 21 refers to arginine vasopressin, invention 22 refers to arginine vasopressin receptor type 11, and invention 23 refers to beta inward rectifier subunit (pancreatic K channel).

12. Claims: 1-9,21-38 (partial); 14 (complete)

INVENTION 24:

A nucleic acid from the gene encoding B2 bradykinin receptor including a polymorphic site according to table 1, column 8, or the complement thereof, an allele-specific oligonucleotide hybridizing to such a polymorphic site, a method of analyzing such a nucleic acid, methods of diagnosing a phenotype related to such a polymorphic site, a computer-readable storage medium as well as a signal carrying data.

13. Claims: 1-9,21-38 (partial)

INVENTION 25 TO INVENTION 64:

A nucleic acid from the genes encoding brain natriuretic protein, bombesin receptor subtype-3, ..., kallistatin, or kallikrein, including a polymorphic site according to table 1, column 8, or the complement thereof, an allele-specific oligonucleotide hybridizing to such a polymorphic site, a method of analyzing such a nucleic acid, methods of diagnosing a phenotype related to such a polymorphic site, a computer-readable storage medium as well as a signal carrying data.

Invention 25 refers to brain natriuretic protein, invention 26 refers to bombesin receptor subtype-3,

invention 63 refers to kallistatin, and invention 64 refers to kallikrein.

14. Claims: 1-9,21-38 (partial); 15 (complete)

INVENTION 65:

A nucleic acid from the gene encoding mineralocorticoid receptor including a polymorphic site according to table 1, column 8, or the complement thereof, an allele-specific



European Patent
Office

LACK OF UNITY OF INVENTION
SHEET B

Application Number
EP 99 25 0150

The Search Division considers that the present European patent application does not comply with the requirements of unity of invention and relates to several inventions or groups of inventions, namely:

oligonucleotide hybridizing to such a polymorphic site, a method of analyzing such a nucleic acid, methods of diagnosing a phenotype related to such a polymorphic site, a computer-readable storage medium as well as a signal carrying data.

15. Claims: 1-9,21-38 (partial); 18 (complete)

INVENTION 66:

A nucleic acid from the gene encoding sodium-calcium exchanger including a polymorphic site according to table 1, column 8, or the complement thereof, an allele-specific oligonucleotide hybridizing to such a polymorphic site, a method of analyzing such a nucleic acid, methods of diagnosing a phenotype related to such a polymorphic site, a computer-readable storage medium as well as a signal carrying data.

16. Claims: 1-9,21-38 (partial)

INVENTION 67 TO INVENTION 75:

A nucleic acid from the genes encoding norepinephrine transporter, neuropeptide Y, ..., preprothyrotropin-releasing hormone, or prostaglandin E receptor EP3 subtype, including a polymorphic site according to table 1, column 8, or the complement thereof, an allele-specific oligonucleotide hybridizing to such a polymorphic site, a method of analyzing such a nucleic acid, methods of diagnosing a phenotype related to such a polymorphic site, a computer-readable storage medium as well as a signal carrying data.

Invention 67 refers to norepinephrine transporter,
invention 68 refers to neuropeptide Y,

invention 74 refers to preprothyrotropin-releasing hormone,
and

invention 75 refers to prostaglandin E receptor EP3 subtype.

17. Claims: 1-9,21-38 (partial); 16 (complete)

INVENTION 76:

A nucleic acid from the gene encoding renin including a polymorphic site according to table 1, column 8, or the complement thereof, an allele-specific oligonucleotide hybridizing to such a polymorphic site, a method of analyzing such a nucleic acid, methods of diagnosing a phenotype related to such a polymorphic site, a computer-readable storage medium as well as a signal

**ANNEX TO THE EUROPEAN SEARCH REPORT
ON EUROPEAN PATENT APPLICATION NO.**

EP 99 25 0150

This annex lists the patent family members relating to the patent documents cited in the above-mentioned European search report.
The members are as contained in the European Patent Office EDP file on
The European Patent Office is in no way liable for these particulars which are merely given for the purpose of information.

14-01-2000

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 9421790 A	29-09-1994	CA 2092455 A	26-09-1994
		AU 6281094 A	11-10-1994
		CA 2158918 A	29-09-1994
		US 5534438 A	09-07-1996
WO 9408048 A	14-04-1994	US 5374525 A	20-12-1994
		AU 670990 B	08-08-1996
		AU 5725094 A	26-04-1994
		CA 2145415 A	14-04-1994
		EP 0665898 A	09-08-1995
		JP 8505043 T	04-06-1996
		NZ 259092 A	28-05-1996
		US 5998145 A	07-12-1999
		US 5589584 A	31-12-1996
		US 5763168 A	09-06-1998
WO 8808457 A	03-11-1988	AU 1728088 A	02-12-1988

EPO FORM P0459

For more details about this annex : see Official Journal of the European Patent Office, No. 12/82



European Patent
Office

LACK OF UNITY OF INVENTION
SHEET B

Application Number
EP 99 25 0150

The Search Division considers that the present European patent application does not comply with the requirements of unity of invention and relates to several inventions or groups of inventions, namely:

carrying data.

18. Claims: 1-9,21-38 (partial); 18 (complete)

INVENTION 77 TO INVENTION 81:

A nucleic acid from the genes encoding SA gene, amiloride-sensitive epithelial sodium channel gamma subunit, ..., thromboxane synthase, or thyrotropin-releasing hormone receptor, including a polymorphic site according to table 1, column 8, or the complement thereof, an allele-specific oligonucleotide hybridizing to such a polymorphic site, a method of analyzing such a nucleic acid, methods of diagnosing a phenotype related to such a polymorphic site, a computer-readable storage medium as well as a signal carrying data.

Invention 77 refers to SA gene,
invention 78 refers to amiloride-sensitive epithelial sodium channel gamma subunit,

invention 80 refers to thromboxane synthase, and
invention 81 refers to thyrotropin-releasing hormone receptor.